US Patent & Trademark Office
Patent Public Search | Text View

WAVEGUIDE COATING OPTIMISATION

Abstract

A method of forming a multi-layer transmission coating for a waveguide, including a step a of determining a first coating parameter and a coating function for each layer to optimise the transmissivity at a plurality of locations along the waveguide for a plurality of different wavelengths. The coating function is chosen from a plurality of allowable coating functions. Next, there is a step b of forming the plurality of layers using the determined coating parameters and coating functions, and a step c of measuring a thickness of at least one layer at each of the plurality of locations. The measurements indicate that the coating function deviates from that selected during the optimisation of step a. Finally, there is a step d of determining a second coating parameter for at least one layer by repeating the optimisation of step a using the coating function derived from the measurements of step c.

Publication Classification

Int. Cl.: **C03C17/00** (20060101); **G02B27/01** (20060101)

U.S. Cl.:

Background/Summary

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of priority of United Kingdom Patent Application no. 2402387.1, filed Feb. 20, 2024, which is hereby incorporated herein by reference in its entirety.

BACKGROUND OF THE DISCLOSURE

Field

[0002] The present disclosure relates to pupil expansion or wavefront replication. More specifically, the present disclosure relates to a waveguide, a display system comprising at least one waveguide according to the disclosure, methods of wavefront replication using the at least one waveguide and a method of waveguiding a wavefront such as a holographic wavefront. Yet more specifically, the present disclosure relates to a method of improving the uniformity of wavefront replicas output by a waveguide and an optical component-such as an optical layer, dielectric layer or dielectric stack—for the same. The present disclosure also relates to a method of forming a transmission coating for a waveguide, a method of designing a transmission coating for a waveguide and a method of fabricating a transmission coating for a waveguide. The transmission coating may be a graded transmission coating.

Technical Background and Introduction

[0003] Light scattered from an object contains both amplitude and phase information. This amplitude and phase information can be captured on, for example, a photosensitive plate by well-known interference techniques to form a holographic recording, or "hologram", comprising interference fringes. The hologram may be reconstructed by illumination with suitable light to form a two-dimensional or three-dimensional holographic reconstruction, or replay image, representative of the original object.

[0004] Computer-generated holography may numerically simulate the interference process. A computer-generated hologram may be calculated by a technique based on a mathematical transformation such as a Fresnel or Fourier transform. These types of holograms may be referred to as Fresnel/Fourier transform holograms or simply Fresnel/Fourier holograms. A Fourier hologram may be considered a Fourier domain/plane representation of the object or a frequency domain/plane representation of the object. A computer-generated hologram may also be calculated by coherent ray tracing or a point cloud technique, for example.

[0005] A computer-generated hologram may be encoded on a spatial light modulator arranged to modulate the amplitude and/or phase of incident light. Light modulation may be achieved using electrically-addressable liquid crystals, optically-addressable liquid crystals or micro-mirrors, for example.

[0006] A spatial light modulator typically comprises a plurality of individually-addressable pixels which may also be referred to as cells or elements. The light modulation scheme may be binary, multilevel or continuous. Alternatively, the device may be continuous (i.e. is not comprised of pixels) and light modulation may therefore be continuous across the device. The spatial light modulator may be reflective meaning that modulated light is output in reflection. The spatial light modulator may equally be transmissive meaning that modulated light is output in transmission.

[0007] A holographic projector may be provided using the system described herein. Such projectors have found application in head-up displays, "HUD".

SUMMARY OF THE DISCLOSURE

[0008] Aspects of the present disclosure are defined in the appended independent claims.

[0009] Broadly, the present disclosure relates to image projection. It relates to a method of image projection and an image projector which comprises a display device. The present disclosure also relates to a projection system comprising the image projector and a viewing system, in which the image projector projects or relays light from the display device to the viewing system. The present disclosure is equally applicable to a monocular and binocular viewing system. The viewing system may comprise a viewer's eye or eyes. The viewing system comprises an optical element having optical power (e.g., lens/es of the human eye) and a viewing plane (e.g., retina of the human eye/s). The projector may be referred to as a 'light engine'. The display device and the image formed (or perceived) using the display device are spatially separated from one another. The image is formed, or perceived by a viewer, on a display plane. In some embodiments, the image is a virtual image and the display plane may be referred to as a virtual image plane. In other embodiments, the image is a real image formed by holographic reconstruction and the image is projected or relayed to the viewing plane. The image is formed by illuminating a diffractive pattern (e.g., hologram) displayed on the display device.

[0010] The display device comprises pixels. The pixels of the display device diffract light. In accordance with well-understood optics, the magnitude of the maximum diffraction angle is determined by the size of the pixels (and other factors such as the wavelength of light).

[0011] In embodiments, the display device is a spatial light modulator such as liquid crystal on silicon ("LCOS") spatial light modulator (SLM). Light propagates over a range of diffraction angles (for example, from zero to the maximum diffractive angle) from the LCOS, towards a viewing entity/system such as a camera or an eye. In some embodiments, magnification techniques may be used to increase the range of available diffraction angles beyond the conventional maximum diffraction angle of an LCOS.

[0012] In some examples, an image (formed from the displayed hologram) is propagated to the eyes. For example, spatially modulated light of an intermediate holographic reconstruction/image formed either in free space or on a screen or other light receiving surface between the display device and the viewer, may be propagated to the viewer.

[0013] In some other examples, the (light of a) hologram itself is propagated to the eyes. For example, spatially modulated light of the hologram (that has not yet been fully transformed to a holographic reconstruction, i.e. image)—that may be informally said to be "encoded" with/by the hologram—is propagated directly to the viewer's eyes. A real or virtual image may be perceived by the viewer. In these embodiments, there is no intermediate holographic reconstruction/image formed between the display device and the viewer. It is sometimes said that, in these embodiments, the lens of the eye performs a hologram-to-image conversion or transform. The projection system, or light engine, may be configured so that the viewer effectively looks directly at the display device.

[0014] In accordance with the principles of well-understood optics, the range of angles of light propagating from a display device that can be viewed, by an eye or other viewing entity/system, varies with the distance between the display device and the viewing entity. At a 1 metre viewing distance, for example, only a small range of angles from an LCOS can propagate through an eye's pupil to form an image at the retina for a given eye position. The range of angles of light rays that are propagated from the display device, which can successfully propagate through an eye's pupil to form an image at the retina for a given eye position, determines the portion of the image that is 'visible' to the viewer. In other words, not all parts of the image are visible from any one point on the viewing plane (e.g., any one eye position within a viewing window such as eye-motion box.)

[0015] In some embodiments, the image perceived by a viewer is a virtual image that appears upstream of the display device—that is, the viewer perceives the image as being further away from them than the display device. Conceptually, it may therefore be considered that the viewer is looking at a virtual image through an 'display device-sized window', which may be very small, for example 1 cm in diameter, at a relatively large distance, e.g., 1 metre. And the user will be viewing

the display device-sized window via the pupil(s) of their eye(s), which can also be very small. Accordingly, the field of view becomes small and the specific angular range that can be seen depends heavily on the eye position, at any given time.

[0016] A pupil expander addresses the problem of how to increase the range of angles of light rays that are propagated from the display device that can successfully propagate through an eye's pupil to form an image. The display device is generally (in relative terms) small and the projection distance is (in relative terms) large. In some embodiments, the projection distance is at least one-such as, at least two-orders of magnitude greater than the diameter, or width, of the entrance pupil and/or aperture of the display device (i.e., size of the array of pixels).

[0017] Use of a pupil expander increases the viewing area (i.e., user's eye-box) laterally, thus enabling some movement of the eye/s to occur, whilst still enabling the user to see the image. As the skilled person will appreciate, in an imaging system, the viewing area (user's eye-box) is the area in which a viewer's eyes can perceive the image. The present disclosure particularly relates to non-infinite virtual image distances—that is, near-field virtual images—but is equally applicable to virtual images formed at infinity or even real images formed downstream of the display device/hologram.

[0018] The display device may have an active or display area having a first dimension that may be less than 10 cms such as less than 5 cms or less than 2 cms. The propagation distance between the display device and viewing system may be greater than 1 m such as greater than 1.5 m or greater than 2 m. The optical propagation distance within the waveguide may be up to 2 m such as up to 1.5 m or up to 1 m. The method may be capable of receiving an image and determining a corresponding hologram of sufficient quality in less than 20 ms such as less than 15 ms or less than 10 ms.

[0019] In some embodiments—described only by way of example of a diffracted or holographic light field or wavefront in accordance with this disclosure—a hologram is configured to route light into a plurality of channels, each channel corresponding to a different part (i.e. sub-area) of an image. The hologram may be represented, such as displayed, on a display device such as a spatial light modulator. When displayed on an appropriate display device, the hologram may spatially modulate light transformable by a viewing system into the image. The channels formed by the diffractive structure are referred to herein as "hologram channels" merely to reflect that they are channels of light encoded by the hologram with image information. It may be said that the light of each channel is in the hologram domain rather than the image or spatial domain. In some embodiments, the hologram is a Fourier or Fourier transform hologram and the hologram domain is therefore the Fourier or frequency domain. The hologram may equally be a Fresnel or Fresnel transform hologram. The hologram is described herein as routing light into a plurality of hologram channels merely to reflect that the image that can be reconstructed from the hologram has a finite size and can be arbitrarily divided into a plurality of image sub-areas, wherein each hologram channel would correspond to each image sub-area. Importantly, the hologram of this example is characterised by how it distributes the image content when illuminated. Specifically, the hologram divides the image content by angle. That is, each point on the image is associated with a unique light ray angle in the spatially modulated light formed by the hologram when illuminated—at least, a unique pair of angles because the hologram is two-dimensional. For the avoidance of doubt, this hologram behaviour is not conventional. The spatially modulated light formed by this special type of hologram, when illuminated, may be arbitrarily divided into a plurality of hologram channels, wherein each hologram channel is defined by a range of light ray angles (in two-dimensions). It will be understood from the foregoing that any hologram channel (i.e. sub-range of light ray angles) that may be considered in the spatially modulated light will be associated with a respective part or sub-area of the image. That is, all the information needed to reconstruct that part or sub-area of the image is contained within a sub-range of angles of the spatially modulated light formed from the hologram of the image. When the spatially modulated light is observed as a whole, there is not

necessarily any evidence of a plurality of discrete light channels. However, in some arrangements, a plurality of spatially separated hologram channels is formed by intentionally leaving areas of the target image, from which the hologram is calculated, blank or empty (i.e., no image content is present).

[0020] Nevertheless, the hologram may still be identified. For example, if only a continuous part or sub-area of the spatially modulated light formed by the hologram is reconstructed, only a sub-area of the image should be visible. If a different, continuous part or sub-area of the spatially modulated light is reconstructed, a different sub-area of the image should be visible. A further identifying feature of this type of hologram is that the shape of the cross-sectional area of any hologram channel substantially corresponds to (i.e. is substantially the same as) the shape of the entrance pupil although the size may be different—at least, at the correct plane for which the hologram was calculated. Each light/hologram channel propagates from the hologram at a different angle or range of angles. Whilst these are example ways of characterising or identifying this type of hologram, other ways may be used. In summary, the hologram disclosed herein is characterised and identifiable by how the image content is distributed within light encoded by the hologram. Again, for the avoidance of any doubt, reference herein to a hologram configured to direct light or angularly-divide an image into a plurality of hologram channels is made by way of example only and the present disclosure is equally applicable to pupil expansion of any type of wavefront or light field including any type of diffractive or diffracted light field.

[0021] Broadly, a system is disclosed herein that provides pupil expansion for an input light field or wavefront, wherein the input light field is, optionally, a diffracted or holographic light field comprising diverging ray bundles. As discussed above, pupil expansion (which may also be referred to as "image replication" or "replication" or "pupil replication") enables the size of the area at/from which a viewer can see an image (or, can receive light of a hologram, which the viewer's eye forms an image) to be increased, by creating one or more replicas of an input light ray (or ray bundle). The pupil expansion can be provided in one or more dimensions. For example, two-dimensional pupil expansion can be provided, with each dimension being substantially orthogonal to the respective other. In embodiments in which the wavefront is a holographic wavefront, the process may be described as hologram replication.

[0022] The system can be provided in a compact and streamlined physical form. This enables the system to be suitable for a broad range of real-world applications, including those for which space is limited and real-estate value is high. For example, it may be implemented in a head-up display (HUD) such as a vehicle or automotive HUD.

[0023] In the present disclosure, the term "replica" is merely used to reflect that spatially modulated light is divided such that a complex light field or wavefront is directed along a plurality of different optical paths. The word "replica" is used to refer to each occurrence or instance of the complex light field or wavefront after a replication event—such as a partial reflection-transmission by a pupil expander. Each replica travels along a different optical path. Some embodiments of the present disclosure relate to propagation of light that is encoded with a hologram, not an image—i.e., light that is spatially modulated with a hologram of an image, not the image itself. The person skilled in the art of holography will appreciate that the complex light field associated with propagation of light encoded with a hologram will change with propagation distance. Use herein of the term "replica" is independent of propagation distance and so the two branches or paths of light associated with a replication event are still referred to as "replicas" of each other even if the branches are a different length, such that the complex light field has evolved differently along each path. That is, two complex light fields are still considered "replicas" in accordance with this disclosure even if they are associated with different propagation distances—providing they have arisen from the same replication event or series of replication events.

[0024] As above, an optical waveguide for use as a pupil expander may lightguide or waveguide a light field or wavefront between a pair of parallel surfaces. This may be achieved by internal

reflection between the parallel surfaces. A first surface of the pair of surfaces may be partially transmissive-reflective. A second surface of the pair of surfaces may be reflective. The light field will therefore be divided at each internal reflection at the first surface such that a plurality of replicas of the light field is transmitted through a region of the first surface that forms an output port of the waveguide. Thus, a viewing window (and eye-box) is enlarged by the waveguide.

[0025] The intensity and spectrum of each successive replica emitted by the waveguide is required to be substantially constant from replica to replica. For example, an integral of the intensity of the light of each replica may be substantially constant, a mean average of the intensity of the light of each replica may be substantially constant and/or a distribution of the intensity of the light of each replica may be substantially constant. Replicas having substantially constant intensity and spectra are referred to herein as being spatially homogenous and a waveguide emitting such replicas is referred to herein as providing spatially homogenous emission. Spatially homogenous replicas advantageously reduce or minimize the variability in the brightness of the image perceived (or different areas thereof) by a viewer moving around the (expanded) viewing window. Furthermore, the overall quality of the hologram received by the viewer (and/or the resulting image perceived by the viewer) may be improved, particularly if the hologram comprises a plurality of different wavelengths of light.

[0026] The intensity of the light being waveguided between the first and second surfaces of the waveguide will decrease after each division of the light field at the first surface. A sub-optimal waveguide may typically comprise a first surface with constant reflectivity. This results in the intensity of each successive replica decreasing. In other words, such conventional waveguides do not provide spatially homogenous emission.

[0027] A graded coating can be applied to the waveguide to provide a first surface having a varying reflectivity and therefore transmissivity. In particular, the coating may be arranged so that the transmissivity of the first surface increases in the direction of waveguiding. The graded coating may be arranged such that the increased transmissivity of the first surface in the direction of waveguiding accounts (e.g. at least partially compensates) for the decrease in the intensity of light being light guided. However, current graded coatings often result in high absorption losses of the waveguided light. Furthermore, these coatings typically allow for very limited control of spectral performance, particularly disadvantageous when the light to be waveguided comprises multiple wavelengths. While some improved graded coatings are available, these are expensive, time consuming and complex to manufacture and generally cannot be manufactured reliably. For example, such coatings may comprise a plurality (often 20 or more) layers of dielectric material, each layer having a unique percentage change of the thickness from a first end of the layer to a second end. Such a complex layered structure may be necessary to provide substantially spatially homogenous emission for the full visible spectrum but is slow to manufacture and difficult to manufacture reliably (for example, a moving grader may be required to be present in the coating chamber).

[0028] The present disclosure addresses a technical problem of providing a spatially homogenous emission from a waveguide pupil expander such that the intensity and spectrum of each replica is substantially similar. The present disclosure proposes an improved waveguide pupil expander which provides substantially homogeneous emission at least at specific wavelengths such as a red, green and blue wavelength. A first surface of the waveguide comprises a plurality of alternating layers of a first and a second dielectric that differ in refractive index. Instead of each layer having a unique percentage change of thickness, in the present disclosure each layer of dielectric has a percentage change of thickness having one of a discrete number of values, the number of discrete values being less than the total number of layers. In some embodiments, the number of discrete or allowable values is no more than four. In some embodiments, the number of layers is at least twice the number of discrete values.

[0029] The multilayer structure resulting from conventional design techniques is very complex and

has low suitability for industrial scale-up. The complexity is partly caused by the need to achieve uniformity at three different wavelengths. This design process has been rethought and focused on restricting it in ways that simplify the manufacturing method that will be needed and relaxing it in other ways that do not impact the image-forming (e.g. holographic) process. In particular, focused has been placed on how the required coating machine may be simplified. If the number of different gradients used in the stack is restricted (e.g. reduced to less than five or even less than four or three) and degrees of freedom are introduced for intermediate wavelengths (e.g. wavelengths between the three hologram wavelengths), adequate optical performance for three-colour hologram replication can still be achieved. The cost of manufacture is significantly reduced because the number of different grading tools is reduced.

[0030] Faster and more reliable methods of manufacture can be used to produce the stack of layers of the present disclosure compared to conventional methods needed to produce a stack of layers where each layer has a unique percentage change of thickness. By limiting the design of the waveguide coating to consider only specific regions of the visible spectrum, such an alternating stack of layers according to the present disclosure can be arranged to provide desirable optical performance such as spatially homogenous emission and low absorption losses at those wavelengths (but not necessarily at other wavelengths). As such, the waveguide of the present disclosure may be particularly suitable for applications in which specific wavelengths of light are to be waveguided rather than a full optical spectrum. For example, in many applications, there may be a need for spatially homogenous emission of red, green and/or blue light from the waveguide, but not other wavelengths (such as in the yellow region of the spectrum). Such applications include head-up displays. The waveguide of the present disclosure may be suitable for such applications. As the skilled person will appreciate, a full colour image can be generated using red, green and blue light.

[0031] According to a first example, a waveguide is provided. The waveguide comprises a pair of parallel/complementary surfaces arranged to provide waveguiding therebetween. A first surface of the pair of parallel surfaces comprises a plurality of layers of a first dielectric and a plurality of layers of second dielectric arranged in an alternating configuration.

[0032] Each layer of the first and second dielectric has a first end and a second end. The first end may be a (light) input end and the second end may be a (light) output end for the final replica or the end of an output port/window. A percentage change in the thickness of each layer from the first end to the second end of that layer has one of a plurality of discrete (e.g. allowable) values. The total number of layers of the first and second dielectric is greater than the total number of discrete values. A difference in refractive index between the first dielectric and second dielectric is greater than 0.4, optionally greater than 0.5.

[0033] References to "percentage change in the thickness" of a layer herein refer to the percentage change in thickness of that layer from the first end of that layer to the second end of that layer, unless specified otherwise. The percentage change in the thickness of a layer may be: $100 \times (t_f - t_i)/t_i$, wherein t_i is the thickness of the respective layer at the first end and t_f us the thickness of the respective layer at the second end.

[0034] Generally, each layer of the plurality of layers may be parallel to the other layers. The first end of each of layer may be aligned with the first end of the other layers and the second end of each layer may be aligned with the other layers. In other words, each layer may have a substantially similar length and/or width as the other layers but may differ in thickness or depth. The plurality of layers of the first and second dielectric may be referred to as a stack of layers or, simply, dielectric stack.

[0035] As used herein, the "alternating configuration" of the layers of the first and second dielectric means that each layer of the first dielectric is separated from the next closest layer(s) of first dielectric by a layer of the second dielectric, and vice versa.

[0036] As used herein, the plurality of discrete values means a plurality of values that are different

to one another. In other words, none of the plurality of discrete values is the same as any other of the discrete values. The magnitude and/or the polarity of each of the discrete values may be different. The discrete values may be referred to as "allowable" or "predetermined" values. The rate of change of the thickness of each layer may be selected during, or determined by, the manufacturing process. For example, if the plurality of layers of the first surfaces of the waveguide are manufactured or fabricated using a shadow mask, the rate of change of the thickness of each layer may be related to the shape of the shadow mask such that each discrete value is associated with a respective shadow mask. The skilled person will appreciate that the "allowable" or "predetermined" discrete values may take any appropriate value that provides desirable optical properties in the context of a particular multilayer structure of first and second dielectrics and that appropriate values could be calculated or determined, for example using simulations. Each discrete value may be positive or negative. A positive values means that the thickness is increasing from the first end to the second end and a negative value means that the thickness is decreasing from the first end to second end.

[0037] Each layer has a percentage change of thickness that is equal to one of the plurality of discrete values. Because there are more layers than there are discrete values, at least two of the layers will have the same percentage change of thickness from the respective first end to the respective second end. In some embodiments, more than two, optionally four or more, optionally 6 or more, optionally 8 or more, optionally 10 or more layers of the plurality of layer may have the same percentage change of thickness from the respective first end to the respective second end (and so these layers may be associated with the same discrete value). In some embodiments, the layers that have the same percentage change of thickness as each other are layers of the same dielectric (i.e. all layers having a particular percentage change of thickness value are layers of the first dielectric or layers of the second dielectric).

[0038] Each of the layers of the first or second dielectric may increase or decrease in thickness in the direction of waveguiding. A first direction may be defined from the first end to the second end of each of the plurality of layers. The pair of parallel surfaces may be arranged to provide waveguiding in the first direction. In other words, light waveguided by the pair of surfaces may interact with the first end before interacting with the second end. Furthermore, each of the layers of the first or second dielectric may increase or decrease in thickness in the first direction.

[0039] The minimum thickness of each of the plurality of layers of the first and second dielectric, as well as the discrete values for the percentage change in thickness, may be selected to provide a waveguide with desired optical properties, in particular with a desired transmissive behaviour such that a spatially homogenous emission is generated. The parameters may depend on (among other things) the material properties (in particular, the refractive index) of the first and second dielectric, the angle of incidence of light entering the waveguide, the wavelength(s) of the incident light and the distance between the pair of parallel surfaces of the waveguide. As will appreciated by the skilled person, there will be a plurality (usually a relatively large number) of arrangements of the plurality of layers that provide a waveguide with desired optical properties. However, common to all of these arrangements are the advantages that the layered structure of the first surface provides desired optical properties and can be manufactured quickly, inexpensively and reliably.

[0040] One example of a fast, inexpensive and reliable method of manufacturing the layers of the first and second dielectric may comprise forming the layers by depositing the respective dielectric material on a waveguide substrate. A shadow mask may be used to control the flow of the dielectric material onto the substrate. Preferably, a shadow mask with a trapezoidal shape may be used. The percentage change in thickness of each layer may be determined by the shape of the mask and the overall thickness of each layer may be determined by the length of time that the dielectric material is allowed to flow. When the mask is trapezoidal, the percentage change in thickness of each layer may correspond to a percentage change in the short base to the long base of the trapezoid shape of the shadow mask. A different mask may be associated with each of the discrete values. For

example, if there are first to fourth percentage change values then four different masks may be used to manufacture the layers of the coating. The waveguide substrate may advantageously simply be moved between different material sources having different masks.

[0041] The plurality of discrete values may comprise or consist of two or more discrete (allowable) values, optionally three or more (allowable) discrete, optionally four or more (allowable) values. The plurality of discrete values may consist of between two and six discrete values, optionally between two and four discrete values, optionally four discrete values. The number of discrete values may preferably be considerably lower than the total number of layers. For example, in some embodiments the total number of layer may be at least 10 layer, optionally at least 15 layers, optionally at least 20 layers.

[0042] The plurality of discrete values may comprise a first value, a second value and a third value. Each layer of the first dielectric may have a percentage change in thickness equal to either the first value or the second value. In other words, there may only be two options for percentage change in the thickness of each layer of the first dielectric.

[0043] At least one layer of the second dielectric may have a percentage change in the thickness equal to the third value. When each of the layers of the first dielectric have either the first discrete value or the second discrete value, this may mean that at least one layer of the second dielectric has a different percentage change in thickness to all of the layers of the first dielectric.

[0044] In addition to the first to third values, the plurality of discrete values may comprise a fourth value. Each layer of the second dielectric may have a percentage change in thickness equal to either the third value or a fourth value. In other words, there may only be two options for the percentage change in the thickness of each layer of the second dielectric. When each of the layers of the first dielectric have either the first discrete value or the second discrete value, this may mean that each layer of the second dielectric has a different percentage thickness to all of the layers of the first dielectric.

[0045] Preferably, the rate of change of the thickness of each layer of the first and second dielectrics may be constant. In other words, each of the layers may have a linear profile. It may be simpler to manufacture dielectric layers having a linear profile. Furthermore, the skilled person will appreciate that it may be more straightforward to calculate/determine discrete allowable values for a stack of layers having a linear profile.

[0046] In some embodiments, the first surface may provide a plurality, n, of light emission zones for light waveguided between the first surface and second surface. The plurality of light emission zones may be distributed along a length of the first surface, in the direction of waveguiding. A replica of the input wavefront may be generated at each emission zone. The first end of each layer of the first and second dielectric layer may be at or adjacent to the first light emission zone. The second end of each layer of the first and second dielectric later may be at or adjacent to the nth light emission zone.

[0047] An angle of internal incidence at each emission zones may be in the range 0 to 70 degrees, preferably in the range of 10 to 50 degrees.

[0048] At other wavelengths (i.e. not the first, second and third visible wavelength), the transmissivity of the first surface may or may not increase in the direction of waveguiding. As previously described, the waveguide of the present disclosure may be particularly advantageous in the context of waveguiding light at discrete wavelengths, rather than waveguiding light having a continuous spectrum. The inventors have appreciated that the parameters such as thickness and the values for the percentage change of thickness of the plurality of layers of the first and second dielectrics may be selected to provide a desired changing transmissivity behaviour at the first, second and third wavelengths only and that there may be no need for the layers to provide that same transmissivity behaviour at other wavelengths. The inventors have found that by confining the problem of providing a changing transmissivity behaviour to only certain wavelengths, a coating comprising multiple layers having the same rate of change of thickness can provide the

desired transmissivity behaviour which, as above, can be manufactured inexpensively, quickly and reliably while still providing acceptable transmissive properties.

[0049] The first wavelength, referred to above, may preferably be in the range 630-670 nm. The second wavelength may preferably be in the range 500-540 nm. The third wavelength may preferably be in the range 430-470 nm. In other words, the first wavelength may correspond to red visible light. The second wavelength may correspond to green visible light. The third wavelength may correspond to blue visible light.

[0050] Preferably, the transmissivity of the first surface at each emission point, $T(n)$, may satisfy the following equation: $T(n)=(T(n-1))/([1-T(n-1)]\times[1-L])$ wherein L is an optical loss factor of the waveguide material.

[0051] The first dielectric may be a first oxide, fluoride, sulfide or nitrate of a first transition metal or semiconductor. The second dielectric may be a second oxide, fluoride, sulfide or nitrate of a second transition metal or semiconductor. In some embodiments, the first dielectric comprises silicon, titanium, tantalum or hafnium. In some embodiments, the second dielectric comprises another of silicon, titanium, tantalum or hafnium.

[0052] Each layer may have a thickness in the range 2 to 300 nm. The thickness of each layer may not fall outside of this range at any point between the first end and second end. In some embodiments, each layer may have a thickness in the range 20 to 300 nm.

[0053] A minimum thickness of each layer may be between 2 and 300 nm, optionally between 20 and 300 nm. A maximum thickness of each layer may be between 2 and 300 nm, optionally between 20 and 300 nm. The minimum thickness of each layer is less than the maximum thickness of that respective layer. The minimum or maximum thickness of each layer may be at the first end of that layer. The other of the minimum or maximum thickness of the respective layer may be at the second end of that layer.

[0054] At least one of the plurality of discrete values for the percentage change in thickness may be positive. At least one of the plurality discrete values of the percentage change in thickness may be negative.

[0055] Each of the plurality of discrete values for the percentage change in thickness may be in the range −150% to +150%.

[0056] The number of layers of the plurality of layers may be at least 10, optionally at least 15, optionally at least 20. The number of layers of the plurality of layers may be in the range 10 to 30, optionally, 15-25.

[0057] According to a first aspect, a method of forming a transmission coating for a waveguide is provided. The transmission coating comprises a plurality of layers. The method comprises a step a of determining a first coating parameter and a coating function for each layer to optimise the transmissivity at a plurality of locations along the waveguide for a plurality of different wavelengths. The coating function is chosen from a plurality of allowable coating functions. Next, there is a step b of forming the plurality of layers using the determined coating parameters and coating functions. Then, there is a step c of measuring a thickness of at least one layer at each of the plurality of locations. The measurements indicate that the coating function deviates from that selected during the optimisation of step a. Finally, there is a step d of determining a second coating parameter for at least one layer by repeating the optimisation of step a using the coating function derived from the measurements of step c. Notably, the coating function is not adjusted. That is, the coating function determined in step a is retained by step d.

[0058] This optimisation improves graded coating performance by adjusting multilayer thicknesses without changing the gradient (mask). It gives more tolerances on the actual thickness gradient (in both linearity and amplitude), thus more tolerances on the geometry of the mask. This reduces time of process development and prototyping by fewer mask iterations. This method can also correct any residual change of masks over time, especially in mass production, to continuously maintain the yield in an automated feedback process.

[0059] In other words, there disclosed herein a method of designing or fabricating a transmission coating for a waveguide. The transmission coating comprises a plurality of layers. The method comprises selecting a (first) coating parameter and (first) coating function, from a plurality of allowable coating function, for each layer to optimise transmissivity at a plurality of locations of the waveguide for a plurality of different wavelengths. The method comprises forming the plurality of layers using the determined coating parameters and coating functions. The method further comprises measuring a thickness of a first layer at the plurality of locations, wherein the measurements indicate that the coating function deviates from that selected during the optimisation. The method is characterized by selecting a different (second) coating parameter for at least one layer (e.g. the first layer and/or different layer/s) by repeating the optimisation using the (second) coating function derived by measurement.

[0060] In reality, however, it is very difficult and time consuming to adjust the mask to exactly the same gradient as designed. In mass production, the gradient may also change from time to time as the mask degrades. In overview, the present disclose describes an optimisation method to help compensate the residual changes of gradient and improve yield and optical performance of the graded coating. More specifically, there is disclosed herein a method to design multilayer achromatic graded transmission coatings using four linear gradients. Each gradient translates to a fixed geometrical shape of the mask of a single dielectric material. In the coating process, the gradient mask physically blocks some of the material from being deposited onto the substrate, thus creating a gradient layer thickness of this material. By carefully combining multiple gradient plus multiple material plus multiple layers, an optimised coating structure can meet our requirement of spatially graded transmission in RGB windows.

[0061] Conventionally, the thickness of a coating may be measured in-situ or after coating to verify the parameters of the coating—particularly the thickness profile. If measurement indicates that a coating mask is no longer providing the desired performance, it is replaced. This is therefore a binary pass-or-fail test and it is accepted in the field that coating masks degrade with time and need replacement (i.e. they periodically "fail"). In industrial or high-volume applications, the process of stopping the production line, replacing a coating mask and then restarting the fabrication processes is costly in time and expense. The inventor recognised that additional tolerance could be achieved by using the pass-or-fail thickness verification measurements to determine an "actual" coating function of the mask. Broadly, the method of the present disclosure comprises identifying a coating function—which is a property or parameter of a physical coating mask—and coating parameter—which is a parameter e.g. coating speed selected by the user—for each layer. The inventor realised that an additional tolerance on the performance of the mask (i.e. a change to the specified coating function) could be achieved by using the degree of freedom associated with the coating parameter to compensate for a degradation of the coating function. In this way, a method is provided that can compensate for wear of the components that are used to apply the coating (such as wear on the apertures of the shadow masks as described above), said wearing causing inaccurate application of the layers producing differing optical properties than those desired. Step a sets an initial set of values that would, under ideal conditions (i.e., no wear on the layer application/coating apparatus), produce the optimised transmissivity along the full length (and, optionally, width) of the waveguide. These values can be produced by simulation, experimental trials or any other suitable method that would be known to the skilled person. This can be referred to as a first optimisation.

[0062] Step b then forms these layers, which can be using the methods described above. If there is no wear on the application apparatus, then there would be no needs to continue the method. However, this is not the case in real world scenarios, which is confirmed in step c as the produced layer thicknesses are compared against the optimised (ideal) values identified in step a.

[0063] A second optimisation (or re-optimisation) is performed. The second optimisation is similar to the first optimisation but uses the coating function identified by measurement rather than the specified/expected coating function of the mask. In other words, the thickness of a (or each) layer is

measured at the plurality of locations and a second (measured) coating function is determined that deviates from the first (expected) coating function. Then, the second optimisation is performed to identify a second coating parameter. In other words, the second coating parameter is a new or updated version of the first coating parameter after the re-optimisation (or second optimisation) has occurred.

[0064] Notably, the inventors have recognised that some degree of imperfection in the shadow mask can be tolerated by performing a second optimisation using a coating function determined by measurement (namely, the pass-or-fail test) and adjusting the selected coating parameter (e.g. coating speed). That is, the inventor surprisingly found that the degree of freedom associated with the coating parameter (e.g. coating rate) can compensate for an imperfection in the coating function (e.g. mask gradient). Conventionally, if a mask were found to provide sub-optimal performance, it would be replaced. However, the inventor has recognised that some degree of sub-optimal performance can be tolerated by utilising a degree of freedom associated with the coating parameter and performing a second (fine-tuning) optimisation using the measured coating function. This approach leads to significant gains in efficiency and extends the lifetime of the mask. In other words, the inventors have surprisingly found that the method yields uniform and stable coating performance even with a shadow mask that has degraded, through use, to be far from its original specification.

[0065] It is possible to extend the method further to create an iterative loop of re-optimisations. In this case, there a further step e of forming or simulating the plurality of layers using the deviated coating function and the adjusted coating parameters of step d. Then, there is a step f of determining or simulating the transmissivity at the plurality of locations along the waveguide for the plurality of different wavelengths. Then, there is a step g of repeating steps c to f until the determined/simulated transmissivity of step f is within a tolerance of the optimised transmissivity of step a. The transmissivity of the produced waveguide is measured in step f. This is to allow the loop, established in step g of steps c to f, to be closed when a waveguide with a sufficiently optimised transmissivity is produced or simulated. The tolerance could be a number of different factors as would be understood by the skilled person, such as each location having a transmissivity within a set percentage error from the optimised transmissivity of step a, or an analysis of a plotted curve of the transmissivity along the length of the waveguide compared to a curve of the optimised transmissivity of step a.

[0066] Each coating function may be a percentage change in thickness of the layer from a first end of the waveguide to a second end of the waveguide along the length of the waveguide. The number of allowable coating functions may be between two and six, optionally between two and four, optionally four. Each determined coating function may be linear. The measurements of step c may indicate that the coating function is non-linear. The layer thickness measurements of step c may achieved by curve fitting measurements of light interacting with the layer to expected behaviours of light with different thicknesses, and thus determining the thickness of the layer. The light measurement may be taken of light reflected by the layer or may use an ellipsometer.

[0067] By having a set number of allowable coating functions, a single application apparatus (the shadow mask, as described above) can apply each layer without needing to be changed. This increases the efficiency of the system and method.

[0068] The first coating parameter may relate to a thickness of the layer at its thinnest point. The formation of the plurality of layers of step b may comprise controlling a flow of a coating material towards the waveguide using a shadow mask. The first coating parameter may be one of: the rate of the flow of coating material towards the waveguide; or a rate at which the waveguide moves through the flow of coating material. Step a may further comprise determining a third coating parameter and step d may further comprise determining a fourth coating parameter, the third and fourth coating parameters being the other of: the rate of the flow of coating material towards the waveguide; or the rate at which the waveguide moves through the flow of coating material. In this

case, the fourth coating parameter is a new or updated version of the third coating parameter after the re-optimisation (or second optimisation) has occurred.

[0069] By adjusting the "base thickness" of the coating of the waveguide, the inventor has found that the transmissivity of each layer can be changed to account for the wear in the application apparatus causing inaccuracies in the coating function of said layers. By using these two parameters, said "base thickness" can be changed without having to change the application apparatus.

[0070] The plurality of layers may comprise at least one dielectric. The formation of the plurality of layers of step b may comprise forming alternate layers of a first material and a second material. The first material may be a first dielectric and the second material may be a second dielectric, the first dielectric being a first oxide, fluoride, sulfide or nitrate of a first transition metal or semiconductor and the second dielectric being a second oxide, fluoride, sulfide or nitrate of a second transition metal or semiconductor. A difference in refractive index between the first material and second material may be greater than 0.4.

[0071] As described above, these materials have a variety of advantageous properties for a waveguide coating.

[0072] The plurality of different wavelengths may comprise a first wavelength, a second wavelength and a third wavelength. The first wavelength may be in the range 630-670 nm, the second wavelength may be in the range 500-540 nm and the third wavelength may be in the range 430-470 nm. The optimised transmissivity at each of the plurality of locations, $T(n)$, along the waveguide may satisfy the following equation: $T(n)=(T(n-1))/([1-T(n-1)]\times[1-L])$ wherein L is an optical loss factor of the waveguide material.

[0073] These follow the optimal properties of the waveguide as described above.

[0074] The optimisations of steps a and e may further comprise accounting for the transmissivity of a further transmission coating on an opposing surface to the transmission coating being applied.

[0075] In this way, the coating being applied can be adjusted to account for inaccuracies in the transmissivity rates of the opposing coated surface that it will be working in parallel with. Alternatively/additionally, the coating being applied can be adjusted to favour or disfavour specific wavelengths of light. This can be to compensate for an optical property of the system or to produce a specific visual effect for the user.

[0076] The deviated coating function may be determined by fitting the original coating function to the measured thickness. This allows the deviated coating function to be accurately identified and quantified for the re-optimisation of step d.

[0077] The measurement of step c may comprise measuring the thickness of each layer. The repeated optimisation of step d may comprise adjusting a coating parameter of each layer at each of the plurality of locations.

[0078] According to a second aspect, a transmission coating for a waveguide applied using the above-described method is provided.

[0079] According to a third aspect, a waveguide comprising a transmission coating applied using the above-described method is provided. The waveguide may further comprise a reflection coating on an opposing surface to the transmission coating.

[0080] According to a fourth aspect, a holographic system is provided comprising a display device arranged to display a hologram of an image and to output spatially modulated light in accordance with the hologram and a first waveguide described above. The holographic system may further comprise a second waveguide.

[0081] According to a fifth aspect, a method of applying a transmission coating to a waveguide is provided. The transmission coating comprises a plurality of layers. The method comprises a step of determining a first coating parameter and selecting a coating function for each layer to optimise the transmissivity at a plurality of locations along the waveguide for a plurality of different wavelengths. The coating function is chosen from a plurality of allowable coating functions such as

four allowable coating functions. Next, there is a step b of forming a first layer using the determined coating parameter and coating function. Then, a step c of identifying that that the coating function deviates from that determined during the optimisation of step a by measuring a thickness of at least the first layer at each of the plurality of locations. Subsequently, there is a step d of repeating the optimisation of the transmissivity at the plurality of locations along the waveguide for the plurality of different wavelengths using the measured coating function of each layer but by adjusting a coating parameter of a second layer. Finally, there is a step e of repeating steps b to d for the second layer and then for each of the subsequent layers.

[0082] This method is similar to that of the first aspect, however in this method the adjustment occurs after each layer is formed, said adjustment taking into account the inaccuracies produced in forming the previous layer. This produces a more optimal coating in a single set of layer applications (i.e., does not require multiple iterations of suboptimal coatings), but produces a less optimal coating compared to the final version made according to the method of the first aspect. This is for at least the reason that viewing the layers holistically allows for better optimisation of the transmissivity across the length of the waveguide.

[0083] The term "hologram" is used to refer to the recording which contains amplitude information or phase information, or some combination thereof, regarding the object. The term "holographic reconstruction" is used to refer to the optical reconstruction of the object which is formed by illuminating the hologram. The system disclosed herein is described as a "holographic projector" because the holographic reconstruction is a real image and spatially-separated from the hologram. The term "replay field" is used to refer to the 2D area within which the holographic reconstruction is formed and fully focused. If the hologram is displayed on a spatial light modulator comprising pixels, the replay field will be repeated in the form of a plurality diffracted orders wherein each diffracted order is a replica of the zeroth-order replay field. The zeroth-order replay field generally corresponds to the preferred or primary replay field because it is the brightest replay field. Unless explicitly stated otherwise, the term "replay field" should be taken as referring to the zeroth-order replay field. The term "replay plane" is used to refer to the plane in space containing all the replay fields. The terms "image", "replay image" and "image region" refer to areas of the replay field illuminated by light of the holographic reconstruction. In some embodiments, the "image" may comprise discrete spots which may be referred to as "image spots" or, for convenience only, "image pixels".

[0084] The terms "encoding", "writing" or "addressing" are used to describe the process of providing the plurality of pixels of the SLM with a respective plurality of control values which respectively determine the modulation level of each pixel. It may be said that the pixels of the SLM are configured to "display" a light modulation distribution in response to receiving the plurality of control values. Thus, the SLM may be said to "display" a hologram and the hologram may be considered an array of light modulation values or levels.

[0085] It has been found that a holographic reconstruction of acceptable quality can be formed from a "hologram" containing only phase information related to the Fourier transform of the original object. Such a holographic recording may be referred to as a phase-only hologram. Embodiments relate to a phase-only hologram but the present disclosure is equally applicable to amplitude-only holography.

[0086] The present disclosure is also equally applicable to forming a holographic reconstruction using amplitude and phase information related to the Fourier transform of the original object. In some embodiments, this is achieved by complex modulation using a so-called fully complex hologram which contains both amplitude and phase information related to the original object. Such a hologram may be referred to as a fully-complex hologram because the value (grey level) assigned to each pixel of the hologram has an amplitude and phase component. The value (grey level) assigned to each pixel may be represented as a complex number having both amplitude and phase components. In some embodiments, a fully-complex computer-generated hologram is calculated.

[0087] Reference may be made to the phase value, phase component, phase information or, simply, phase of pixels of the computer-generated hologram or the spatial light modulator as shorthand for "phase-delay". That is, any phase value described is, in fact, a number (e.g. in the range 0 to 2π) which represents the amount of phase retardation provided by that pixel. For example, a pixel of the spatial light modulator described as having a phase value of $\pi/2$ will retard the phase of received light by $\pi/2$ radians. In some embodiments, each pixel of the spatial light modulator is operable in one of a plurality of possible modulation values (e.g. phase delay values). The term "grey level" may be used to refer to the plurality of available modulation levels. For example, the term "grey level" may be used for convenience to refer to the plurality of available phase levels in a phase-only modulator even though different phase levels do not provide different shades of grey. The term "grey level" may also be used for convenience to refer to the plurality of available complex modulation levels in a complex modulator.

[0088] The hologram therefore comprises an array of grey levels—that is, an array of light modulation values such as an array of phase-delay values or complex modulation values. The hologram is also considered a diffractive pattern because it is a pattern that causes diffraction when displayed on a spatial light modulator and illuminated with light having a wavelength comparable to, generally less than, the pixel pitch of the spatial light modulator. Reference is made herein to combining the hologram with other diffractive patterns such as diffractive patterns functioning as a lens or grating. For example, a diffractive pattern functioning as a grating may be combined with a hologram to translate the replay field on the replay plane or a diffractive pattern functioning as a lens may be combined with a hologram to focus the holographic reconstruction on a replay plane in the near field.

[0089] Although different embodiments and groups of embodiments may be disclosed separately in the detailed description which follows, any feature of any embodiment or group of embodiments may be combined with any other feature or combination of features of any embodiment or group of embodiments. That is, all possible combinations and permutations of features disclosed in the present disclosure are envisaged.

Description

BRIEF DESCRIPTION OF THE DRAWINGS

[0090] Specific embodiments are described by way of example only with reference to the following figures:

[0091] FIG. **1** is a schematic showing a reflective SLM producing a holographic reconstruction on a screen;

[0092] FIG. **2**A illustrates a first iteration of an example Gerchberg-Saxton type algorithm;

[0093] FIG. **2**B illustrates the second and subsequent iterations of the example Gerchberg-Saxton type algorithm;

[0094] FIG. **2**C illustrates alternative second and subsequent iterations of the example Gerchberg-Saxton type algorithm;

[0095] FIG. **3** is a schematic of a reflective LCOS SLM;

[0096] FIG. **4**A shows an image comprising a plurality of image areas (bottom) and corresponding hologram comprising a plurality of hologram components (top);

[0097] FIG. **4**B shows a hologram characterised by the routing or channelling of holographically encoded light into a plurality of discrete hologram channels;

[0098] FIG. **5** shows a system arranged to route the light content of each hologram channel of FIG. **4**B through a different optical path to the eye;

[0099] FIG. **6** shows a perspective view of a pair of stacked image replicators arranged for expanding a beam in two dimensions;

[0100] FIG. **7** shows a cross-sectional schematic view of a first waveguide according to the present disclosure;

[0101] FIG. **8** shows a close-up cross-sectional schematic view of a portion of the first waveguide of FIG. **7**;

[0102] FIG. **9** shows a graph of the ideal increasing transmissivity of a waveguide in the direction of waveguiding;

[0103] FIG. **10**A, FIG. **10**B and FIG. **10**C show graphs related to a First Example of a waveguide according to the present disclosure comprising 12 alternating layers of SiO.sub.2 and TiO.sub.2 in which FIGS. **10**A and **10**B respectively show the thickness of the SiO.sub.2 and TiO.sub.2 layers and FIG. **10**C shows the transmissivity of a first surface of the waveguide of the First Example compared to the ideal transmissivity at red, green and blue wavelengths for which the waveguide has been designed;

[0104] FIG. **11** shows a graph of the transmissivity of the waveguide of the First Example at other wavelengths, compared to the ideal transmissivity;

[0105] FIG. **12** shows a cross-sectional schematic of a portion of an apparatus for manufacturing a waveguide according to the present disclosure in which a waveguide substrate is passing under a source of dielectric material;

[0106] FIG. **13** is a schematic view of a shadow mask of the apparatus of FIG. **12** comprising four apertures, the cross-section being in a plane that is orthogonal to the plane of the cross-section of FIG. **12**;

[0107] FIG. **14**A, FIG. **14**B, FIG. **14**C and FIG. **14**D show cross-sectional schematic views of four different waveguide substrates each having a layer of dielectric material formed on a first surface using the mask of FIG. **13**, wherein each layer has been formed using a different aperture of the mask;

[0108] FIG. **15** illustrates a first method of applying a transmission coating to a waveguide according to the present disclosure;

[0109] FIG. **16** illustrates an extended version of the first method of applying a transmission coating to a waveguide according to the present disclosure;

[0110] FIG. **17** illustrates a second method of applying a transmission coating to a waveguide according to the present disclosure;

[0111] FIG. **18**A, FIG. **18**B and FIG. **18**C show graphs related to an example of a simulated optimised transmission coating for a waveguide according to the present disclosure comprising 12 alternating layers of a first and second dielectric material in which FIGS. **18**A and **18**B respectively show the thickness of the first and second dielectric material layers and FIG. **18**C shows the transmissivity of the simulated optimised transmission coating compared to the ideal transmissivity at red, green and blue wavelengths for which the transmission coating has been designed; and

[0112] FIG. **19**A, FIG. **19**B and FIG. **19**C show graphs related to a first example of a transmission coating for a waveguide that has been applied by a method according to the present disclosure comprising 12 alternating layers of the first and second dielectric material in which FIGS. **19**A and **19**B respectively show the thickness of the first and second dielectric material layers and FIG. **19**C shows the transmissivity of the transmission coating compared to the ideal transmissivity at red, green and blue wavelengths for which the transmission coating has been designed.

[0113] The same reference numbers will be used throughout the drawings to refer to the same or like parts.

DETAILED DESCRIPTION OF EMBODIMENTS

[0114] The present invention is not restricted to the embodiments described in the following but extends to the full scope of the appended claims. That is, the present invention may be embodied in different forms and should not be construed as limited to the described embodiments, which are set out for the purpose of illustration.

[0115] Terms of a singular form may include plural forms unless specified otherwise.

[0116] A structure described as being formed at an upper portion/lower portion of another structure or on/under the other structure should be construed as including a case where the structures contact each other and, moreover, a case where a third structure is disposed there between.

[0117] In describing a time relationship—for example, when the temporal order of events is described as "after", "subsequent", "next", "before" or suchlike—the present disclosure should be taken to include continuous and non-continuous events unless otherwise specified. For example, the description should be taken to include a case which is not continuous unless wording such as "just", "immediate" or "direct" is used.

[0118] Although the terms "first", "second", etc. may be used herein to describe various elements, these elements are not to be limited by these terms. These terms are only used to distinguish one element from another. For example, a first element could be termed a second element, and, similarly, a second element could be termed a first element, without departing from the scope of the appended claims.

[0119] Features of different embodiments may be partially or overall coupled to or combined with each other, and may be variously inter-operated with each other. Some embodiments may be carried out independently from each other, or may be carried out together in co-dependent relationship.

Optical Configuration

[0120] FIG. **1** shows an embodiment in which a computer-generated hologram is encoded on a single spatial light modulator. The computer-generated hologram is a Fourier transform of the object for reconstruction. It may therefore be said that the hologram is a Fourier domain or frequency domain or spectral domain representation of the object. In this embodiment, the spatial light modulator is a reflective liquid crystal on silicon, "LCOS", device. The hologram is encoded on the spatial light modulator and a holographic reconstruction is formed at a replay field, for example, a light receiving surface such as a screen or diffuser.

[0121] A light source **110**, for example a laser or laser diode, is disposed to illuminate the SLM **140** via a collimating lens **111**. The collimating lens causes a generally planar wavefront of light to be incident on the SLM. In FIG. **1**, the direction of the wavefront is off-normal (e.g. two or three degrees away from being truly orthogonal to the plane of the transparent layer). However, in other embodiments, the generally planar wavefront is provided at normal incidence and a beam splitter arrangement is used to separate the input and output optical paths. In the embodiment shown in FIG. **1**, the arrangement is such that light from the light source is reflected off a mirrored rear surface of the SLM and interacts with a light-modulating layer to form an exit wavefront **112**. The exit wavefront **112** is applied to optics including a Fourier transform lens **120**, having its focus at a screen **125**. More specifically, the Fourier transform lens **120** receives a beam of modulated light from the SLM **140** and performs a frequency-space transformation to produce a holographic reconstruction at the screen **125**.

[0122] Notably, in this type of holography, each pixel of the hologram contributes to the whole reconstruction. There is not a one-to-one correlation between specific points (or image pixels) on the replay field and specific light-modulating elements (or hologram pixels). In other words, modulated light exiting the light-modulating layer is distributed across the replay field.

[0123] In these embodiments, the position of the holographic reconstruction in space is determined by the dioptric (focusing) power of the Fourier transform lens. In the embodiment shown in FIG. **1**, the Fourier transform lens is a physical lens. That is, the Fourier transform lens is an optical Fourier transform lens and the Fourier transform is performed optically. Any lens can act as a Fourier transform lens but the performance of the lens will limit the accuracy of the Fourier transform it performs. The skilled person understands how to use a lens to perform an optical Fourier transform.

Hologram Calculation

[0124] In some embodiments, the computer-generated hologram is a Fourier transform hologram, or simply a Fourier hologram or Fourier-based hologram, in which an image is reconstructed in the

far field by utilising the Fourier transforming properties of a positive lens. The Fourier hologram is calculated by Fourier transforming the desired light field in the replay plane back to the lens plane. Computer-generated Fourier holograms may be calculated using Fourier transforms.

[0125] A Fourier transform hologram may be calculated using an algorithm such as the Gerchberg-Saxton algorithm. Furthermore, the Gerchberg-Saxton algorithm may be used to calculate a hologram in the Fourier domain (i.e. a Fourier transform hologram) from amplitude-only information in the spatial domain (such as a photograph). The phase information related to the object is effectively "retrieved" from the amplitude-only information in the spatial domain. In some embodiments, a computer-generated hologram is calculated from amplitude-only information using the Gerchberg-Saxton algorithm or a variation thereof.

[0126] The Gerchberg Saxton algorithm considers the situation when intensity cross-sections of a light beam, I.sub.A(x, y) and I.sub.B(x, y), in the planes A and B respectively, are known and I.sub.A(x, y) and I.sub.B(x, y) are related by a single Fourier transform. With the given intensity cross-sections, an approximation to the phase distribution in the planes A and B, Ψ.sub.A(x, y) and Ψ.sub.B(x, y) respectively, is found. The Gerchberg-Saxton algorithm finds solutions to this problem by following an iterative process. More specifically, the Gerchberg-Saxton algorithm iteratively applies spatial and spectral constraints while repeatedly transferring a data set (amplitude and phase), representative of I.sub.A(x, y) and I.sub.B(x, y), between the spatial domain and the Fourier (spectral or frequency) domain. The corresponding computer-generated hologram in the spectral domain is obtained through at least one iteration of the algorithm. The algorithm is convergent and arranged to produce a hologram representing an input image. The hologram may be an amplitude-only hologram, a phase-only hologram or a fully complex hologram.

[0127] In some embodiments, a phase-only hologram is calculated using an algorithm based on the Gerchberg-Saxton algorithm such as described in British patent 2,498,170 or 2,501,112 which are hereby incorporated in their entirety by reference. However, embodiments disclosed herein describe calculating a phase-only hologram by way of example only. In these embodiments, the Gerchberg-Saxton algorithm retrieves the phase information Ψ[u, v] of the Fourier transform of the data set which gives rise to a known amplitude information T[x, y], wherein the amplitude information T[x, y] is representative of a target image (e.g. a photograph). Since the magnitude and phase are intrinsically combined in the Fourier transform, the transformed magnitude and phase contain useful information about the accuracy of the calculated data set. Thus, the algorithm may be used iteratively with feedback on both the amplitude and the phase information. However, in these embodiments, only the phase information Ψ[u, v] is used as the hologram to form a holographic representative of the target image at an image plane. The hologram is a data set (e.g. 2D array) of phase values.

[0128] In other embodiments, an algorithm based on the Gerchberg-Saxton algorithm is used to calculate a fully-complex hologram. A fully-complex hologram is a hologram having a magnitude component and a phase component. The hologram is a data set (e.g. 2D array) comprising an array of complex data values wherein each complex data value comprises a magnitude component and a phase component.

[0129] In some embodiments, the algorithm processes complex data and the Fourier transforms are complex Fourier transforms. Complex data may be considered as comprising (i) a real component and an imaginary component or (ii) a magnitude component and a phase component. In some embodiments, the two components of the complex data are processed differently at various stages of the algorithm.

[0130] FIG. **2**A illustrates the first iteration of an algorithm in accordance with some embodiments for calculating a phase-only hologram. The input to the algorithm is an input image **210** comprising a 2D array of pixels or data values, wherein each pixel or data value is a magnitude, or amplitude, value. That is, each pixel or data value of the input image **210** does not have a phase component. The input image **210** may therefore be considered a magnitude-only or amplitude-only or intensity-

only distribution. An example of such an input image **210** is a photograph or one frame of video comprising a temporal sequence of frames. The first iteration of the algorithm starts with a data forming step **202**A comprising assigning a random phase value to each pixel of the input image, using a random phase distribution (or random phase seed) **230**, to form a starting complex data set wherein each data element of the set comprising magnitude and phase. It may be said that the starting complex data set is representative of the input image in the spatial domain.

[0131] First processing block **250** receives the starting complex data set and performs a complex Fourier transform to form a Fourier transformed complex data set. Second processing block **253** receives the Fourier transformed complex data set and outputs a hologram **280**A. In some embodiments, the hologram **280**A is a phase-only hologram. In these embodiments, second processing block **253** quantises each phase value and sets each amplitude value to unity in order to form hologram **280**A. Each phase value is quantised in accordance with the phase-levels which may be represented on the pixels of the spatial light modulator which will be used to "display" the phase-only hologram. For example, if each pixel of the spatial light modulator provides **256** different phase levels, each phase value of the hologram is quantised into one phase level of the **256** possible phase levels. Hologram **280**A is a phase-only Fourier hologram which is representative of an input image. In other embodiments, the hologram **280**A is a fully complex hologram comprising an array of complex data values (each including an amplitude component and a phase component) derived from the received Fourier transformed complex data set. In some embodiments, second processing block **253** constrains each complex data value to one of a plurality of allowable complex modulation levels to form hologram **280**A. The step of constraining may include setting each complex data value to the nearest allowable complex modulation level in the complex plane. It may be said that hologram **280**A is representative of the input image in the spectral or Fourier or frequency domain. In some embodiments, the algorithm stops at this point.

[0132] However, in other embodiments, the algorithm continues as represented by the dotted arrow in FIG. **2**A. In other words, the steps which follow the dotted arrow in FIG. **2**A are optional (i.e. not essential to all embodiments).

[0133] Third processing block **256** receives the modified complex data set from the second processing block **253** and performs an inverse Fourier transform to form an inverse Fourier transformed complex data set. It may be said that the inverse Fourier transformed complex data set is representative of the input image in the spatial domain.

[0134] Fourth processing block **259** receives the inverse Fourier transformed complex data set and extracts the distribution of magnitude values **211**A and the distribution of phase values **213**A. Optionally, the fourth processing block **259** assesses the distribution of magnitude values **211**A. Specifically, the fourth processing block **259** may compare the distribution of magnitude values **211**A of the inverse Fourier transformed complex data set with the input image **510** which is itself, of course, a distribution of magnitude values. If the difference between the distribution of magnitude values **211**A and the input image **210** is sufficiently small, the fourth processing block **259** may determine that the hologram **280**A is acceptable. That is, if the difference between the distribution of magnitude values **211**A and the input image **210** is sufficiently small, the fourth processing block **259** may determine that the hologram **280**A is a sufficiently-accurate representative of the input image **210**. In some embodiments, the distribution of phase values **213**A of the inverse Fourier transformed complex data set is ignored for the purpose of the comparison. It will be appreciated that any number of different methods for comparing the distribution of magnitude values **211**A and the input image **210** may be employed and the present disclosure is not limited to any particular method. In some embodiments, a mean square difference is calculated and if the mean square difference is less than a threshold value, the hologram **280**A is deemed acceptable. If the fourth processing block **259** determines that the hologram **280**A is not acceptable, a further iteration of the algorithm may be performed. However, this comparison step is not essential and in other embodiments, the number of iterations of the algorithm performed is

predetermined or preset or user-defined.

[0135] FIG. **2**B represents a second iteration of the algorithm and any further iterations of the algorithm. The distribution of phase values **213**A of the preceding iteration is fed-back through the processing blocks of the algorithm. The distribution of magnitude values **211**A is rejected in favour of the distribution of magnitude values of the input image **210**. In the first iteration, the data forming step **202**A formed the first complex data set by combining distribution of magnitude values of the input image **210** with a random phase distribution **230**. However, in the second and subsequent iterations, the data forming step **202**B comprises forming a complex data set by combining (i) the distribution of phase values **213**A from the previous iteration of the algorithm with (ii) the distribution of magnitude values of the input image **210**.

[0136] The complex data set formed by the data forming step **202**B of FIG. **2**B is then processed in the same way described with reference to FIG. **2**A to form second iteration hologram **280**B. The explanation of the process is not therefore repeated here. The algorithm may stop when the second iteration hologram **280**B has been calculated. However, any number of further iterations of the algorithm may be performed. It will be understood that the third processing block **256** is only required if the fourth processing block **259** is required or a further iteration is required. The output hologram **280**B generally gets better with each iteration. However, in practice, a point is usually reached at which no measurable improvement is observed or the positive benefit of performing a further iteration is out-weighted by the negative effect of additional processing time. Hence, the algorithm is described as iterative and convergent.

[0137] FIG. **2**C represents an alternative embodiment of the second and subsequent iterations. The distribution of phase values **213**A of the preceding iteration is fed-back through the processing blocks of the algorithm. The distribution of magnitude values **211**A is rejected in favour of an alternative distribution of magnitude values. In this alternative embodiment, the alternative distribution of magnitude values is derived from the distribution of magnitude values **211** of the previous iteration. Specifically, processing block **258** subtracts the distribution of magnitude values of the input image **210** from the distribution of magnitude values **211** of the previous iteration, scales that difference by a gain factor α and subtracts the scaled difference from the input image **210**. This is expressed mathematically by the following equations, wherein the subscript text and numbers indicate the iteration number:

[00001] $R_{n+1}[x,y] = F'\{\exp(i\ \psi_n[u,v])\}$

$\psi_n[u,v] = \angle F\{\ \eta\ .Math.\ \exp(i\angle R_n[x,y])\}$

$\eta = T[x,y] - \alpha\ (\ .Math.\ R_n[x,y]\ .Math.\ -T[x,y])$

where: [0138] F' is the inverse Fourier transform; [0139] F is the forward Fourier transform; [0140] R[x, y] is the complex data set output by the third processing block **256**; [0141] T[x, y] is the input or target image; [0142] $\angle$ is the phase component; [0143] Ψ is the phase-only hologram **280**B; [0144] η is the new distribution of magnitude values **211**B; and [0145] α is the gain factor.

[0146] The gain factor α may be fixed or variable. In some embodiments, the gain factor α is determined based on the size and rate of the incoming target image data. In some embodiments, the gain factor α is dependent on the iteration number. In some embodiments, the gain factor α is solely function of the iteration number.

[0147] The embodiment of FIG. **2**C is the same as that of FIG. **2**A and FIG. **2**B in all other respects. It may be said that the phase-only hologram $\Psi(u, v)$ comprises a phase distribution in the frequency or Fourier domain.

[0148] In some embodiments, the Fourier transform is performed using the spatial light modulator. Specifically, the hologram data is combined with second data providing optical power. That is, the data written to the spatial light modulation comprises hologram data representing the object and lens data representative of a lens. When displayed on a spatial light modulator and illuminated with light, the lens data emulates a physical lens—that is, it brings light to a focus in the same way as the corresponding physical optic. The lens data therefore provides optical, or focusing, power. In these embodiments, the physical Fourier transform lens **120** of FIG. **1** may be omitted. It is known

how to calculate data representative of a lens. The data representative of a lens may be referred to as a software lens. For example, a phase-only lens may be formed by calculating the phase delay caused by each point of the lens owing to its refractive index and spatially-variant optical path length. For example, the optical path length at the centre of a convex lens is greater than the optical path length at the edges of the lens. An amplitude-only lens may be formed by a Fresnel zone plate. It is also known in the art of computer-generated holography how to combine data representative of a lens with a hologram so that a Fourier transform of the hologram can be performed without the need for a physical Fourier lens. In some embodiments, lensing data is combined with the hologram by simple addition such as simple vector addition. In some embodiments, a physical lens is used in conjunction with a software lens to perform the Fourier transform. Alternatively, in other embodiments, the Fourier transform lens is omitted altogether such that the holographic reconstruction takes place in the far-field. In further embodiments, the hologram may be combined in the same way with grating data—that is, data arranged to perform the function of a grating such as image steering. Again, it is known in the field how to calculate such data. For example, a phase-only grating may be formed by modelling the phase delay caused by each point on the surface of a blazed grating. An amplitude-only grating may be simply superimposed with an amplitude-only hologram to provide angular steering of the holographic reconstruction. The second data providing lensing and/or steering may be referred to as a light processing function or light processing pattern to distinguish from the hologram data which may be referred to as an image forming function or image forming pattern.

[0149] In some embodiments, the Fourier transform is performed jointly by a physical Fourier transform lens and a software lens. That is, some optical power which contributes to the Fourier transform is provided by a software lens and the rest of the optical power which contributes to the Fourier transform is provided by a physical optic or optics.

[0150] In some embodiments, there is provided a real-time engine arranged to receive image data and calculate holograms in real-time using the algorithm. In some embodiments, the image data is a video comprising a sequence of image frames. In other embodiments, the holograms are pre-calculated, stored in computer memory and recalled as needed for display on a SLM. That is, in some embodiments, there is provided a repository of predetermined holograms.

[0151] Embodiments relate to Fourier holography and Gerchberg-Saxton type algorithms by way of example only. The present disclosure is equally applicable to Fresnel holography and Fresnel holograms which may be calculated by a similar method. The present disclosure is also applicable to holograms calculated by other techniques such as those based on point cloud methods.

Light Modulation

[0152] A spatial light modulator may be used to display the diffractive pattern including the computer-generated hologram. If the hologram is a phase-only hologram, a spatial light modulator which modulates phase is required. If the hologram is a fully-complex hologram, a spatial light modulator which modulates phase and amplitude may be used or a first spatial light modulator which modulates phase and a second spatial light modulator which modulates amplitude may be used.

[0153] In some embodiments, the light-modulating elements (i.e. the pixels) of the spatial light modulator are cells containing liquid crystal. That is, in some embodiments, the spatial light modulator is a liquid crystal device in which the optically-active component is the liquid crystal. Each liquid crystal cell is configured to selectively-provide a plurality of light modulation levels. That is, each liquid crystal cell is configured at any one time to operate at one light modulation level selected from a plurality of possible light modulation levels. Each liquid crystal cell is dynamically-reconfigurable to a different light modulation level from the plurality of light modulation levels. In some embodiments, the spatial light modulator is a reflective liquid crystal on silicon (LCOS) spatial light modulator but the present disclosure is not restricted to this type of spatial light modulator.

[0154] A LCOS device provides a dense array of light modulating elements, or pixels, within a small aperture (e.g. a few centimetres in width). The pixels are typically approximately 10 microns or less which results in a diffraction angle of a few degrees meaning that the optical system can be compact. It is easier to adequately illuminate the small aperture of a LCOS SLM than it is the larger aperture of other liquid crystal devices. An LCOS device is typically reflective which means that the circuitry which drives the pixels of a LCOS SLM can be buried under the reflective surface. The results in a higher aperture ratio. In other words, the pixels are closely packed meaning there is very little dead space between the pixels. This is advantageous because it reduces the optical noise in the replay field. A LCOS SLM uses a silicon backplane which has the advantage that the pixels are optically flat. This is particularly important for a phase modulating device.

[0155] A suitable LCOS SLM is described below, by way of example only, with reference to FIG. 3. An LCOS device is formed using a single crystal silicon substrate 302. It has a 2D array of square planar aluminium electrodes 301, spaced apart by a gap 301*a*, arranged on the upper surface of the substrate. Each of the electrodes 301 can be addressed via circuitry 302*a* buried in the substrate 302. Each of the electrodes forms a respective planar mirror. An alignment layer 303 is disposed on the array of electrodes, and a liquid crystal layer 304 is disposed on the alignment layer 303. A second alignment layer 305 is disposed on the planar transparent layer 306, e.g. of glass. A single transparent electrode 307 e.g. of ITO is disposed between the transparent layer 306 and the second alignment layer 305.

[0156] Each of the square electrodes 301 defines, together with the overlying region of the transparent electrode 307 and the intervening liquid crystal material, a controllable phase-modulating element 308, often referred to as a pixel. The effective pixel area, or fill factor, is the percentage of the total pixel which is optically active, taking into account the space between pixels 301*a*. By control of the voltage applied to each electrode 301 with respect to the transparent electrode 307, the properties of the liquid crystal material of the respective phase modulating element may be varied, thereby to provide a variable delay to light incident thereon. The effect is to provide phase-only modulation to the wavefront, i.e. no amplitude effect occurs.

[0157] The described LCOS SLM outputs spatially modulated light in reflection. Reflective LCOS SLMs have the advantage that the signal lines, gate lines and transistors are below the mirrored surface, which results in high fill factors (typically greater than 90%) and high resolutions. Another advantage of using a reflective LCOS spatial light modulator is that the liquid crystal layer can be half the thickness than would be necessary if a transmissive device were used. This greatly improves the switching speed of the liquid crystal (a key advantage for the projection of moving video images). However, the teachings of the present disclosure may equally be implemented using a transmissive LCOS SLM.

Light Channelling

[0158] The optical system disclosed herein is applicable to pupil expansion with any diffracted light field. In some embodiments, the diffracted light field is a holographic light field—that is, a complex light field that has been spatially modulated in accordance with a hologram of an image, not the image itself. In some embodiments, the hologram is a special type of hologram that angularly divides/channels the image content. This type of hologram is described further herein merely as an example of a diffracted light field that is compatible with the present disclosure. Other types of hologram may be used in conjunction with the display systems and light engines disclosed herein.

[0159] A display system and method is described herebelow, which comprises a waveguide pupil expander. As will be familiar to the skilled reader, a waveguide may be regarded as being a 'pupil expander' because it can be used to increase the area over (or, within) which the light emitted by a relatively small light emitter—such as a relatively small SLM or other pixelated display device as used in the arrangements described herein—can be viewed by a human viewer or other viewing

system that is located at a distance, such as a relatively large distance, away from the light emitter. The waveguide achieves this by increasing the number of transmission points from which the light is output, towards the viewer. As a result, the light may be seen from a plurality of different viewer locations and, for example, the viewer may be able to move their head, and therefore their line of sight, whilst still being able to see the light from the light emitter. Thus, it can be said that the viewer's 'eye-box' or 'eye-motion box' is enlarged, through use of a waveguide pupil expander. This has many useful applications, for example but not limited to head-up displays, for example but not limited to automotive head-up displays.

[0160] A display system as described herein may be configured to guide light, such as a diffracted light field, through a waveguide pupil expander in order to provide pupil expansion in at least one dimension, for example in two dimensions. The diffracted light field may comprise light output by a spatial light modulator (SLM), such as an LCOS SLM. For example, that diffracted light field may comprise light that is encoded by a hologram displayed by the SLM. For example, that diffracted light field may comprise light of a holographically reconstructed image, corresponding to a hologram displayed by the SL M. The hologram may comprise a computer-generated hologram (CGH) such as, but not limited to, a point-cloud hologram, a Fresnel hologram, or a Fourier hologram. The hologram may be referred to as being a 'diffractive structure' or a 'modulation pattern'. The SLM or other display device may be arranged to display a diffractive pattern (or, modulation pattern) that comprises the hologram and one or more other elements such as a software lens or diffraction grating, in a manner that will be familiar to the skilled reader.

[0161] The hologram may be calculated to provide channelling of the diffracted light field. This is described in detail in each of GB2101666.2 (published as GB2603517A), GB2101667.0 (published as GB2603518A), and GB2112213.0 (published as GB2610203A), each of which is hereby incorporated by reference herein in its entirety. In general terms, the hologram may be calculated to correspond to an image that is to be holographically reconstructed. That image, to which the hologram corresponds, may be referred to as an 'input image' or a 'target image'. The hologram may be calculated so that, when it is displayed on an SLM and suitably illuminated, it forms a light field (output by the SLM) that comprises a cone of spatially modulated light. In some embodiments the cone comprises a plurality of continuous light channels of spatially modulated light that correspond with respective continuous regions of the image. However, the present disclosure is not limited to a hologram of this type.

[0162] Although we refer to a 'hologram' or to a 'computer-generated hologram (CGH)' herein, it will be appreciated that an SLM may be configured to dynamically display a plurality of different holograms in succession or according to a sequence. The systems and methods described herein are applicable to the dynamic display of a plurality of different holograms.

[0163] FIGS. **4**A, **4**B, **4**C and **5** show an example of a type of hologram that may be displayed on a display device such as an SLM, which can be used in conjunction with a pupil expander as disclosed herein. However, this example should not be regarded as limiting with respect to the present disclosure.

[0164] FIG. **4**A shows an image **452** for projection comprising eight image areas/components, V1 to V8. FIG. **4**A shows eight image components by way of example only and the image **452** may be divided into any number of components. FIG. **4**A also shows an encoded light pattern **454** (i.e., hologram) that can reconstruct the image **452**—e.g., when transformed by the lens of a suitable viewing system. The encoded light pattern **454** comprises first to eighth sub-holograms or components, H1 to H8, corresponding to the first to eighth image components/areas, V1 to V8. FIG. **4**A further shows how a hologram may decompose the image content by angle. The hologram may therefore be characterised by the channelling of light that it performs. This is illustrated in FIG. **4**B. Specifically, the hologram in this example directs light into a plurality of discrete areas. The discrete areas are discs in the example shown but other shapes are envisaged. The size and shape of the optimum disc may, after propagation through the waveguide, be related to the size and

shape of the entrance pupil of the viewing system.

[0165] FIG. **5** shows a viewing system **500**, including a display device that displays a hologram that has been calculated as illustrated in FIGS. **4**A and **4**B.

[0166] The viewing system **500** comprises a display device, which in this arrangement comprises an LCOS **502**. The LCOS **502** is arranged to display a modulation pattern (or 'diffractive pattern') comprising the hologram and to project light that has been holographically encoded towards an eye **505** that comprises a pupil that acts as an aperture **504**, a lens **509**, and a retina (not shown) that acts as a viewing plane. There is a light source (not shown) arranged to illuminate the LCOS **502**. The lens **509** of the eye **505** performs a hologram-to-image transformation. The light source may be of any suitable type. For example, it may comprise a laser light source.

[0167] The viewing system **500** further comprises a waveguide **508** positioned between the LCOS **502** and the eye **505**. The presence of the waveguide **508** enables all angular content from the LCOS **502** to be received by the eye, even at the relatively large projection distance shown. This is because the waveguide **508** acts as a pupil expander, in a manner that is well known and so is described only briefly herein.

[0168] In brief, the waveguide **508** shown in FIG. **5** comprises a substantially elongate formation. In this example, the waveguide **508** comprises an optical slab of refractive material, but other types of waveguide are also well known and may be used. The waveguide **508** is located so as to intersect the light cone (i.e., the diffracted light field) that is projected from the LCOS **502**, for example at an oblique angle. In this example, the size, location, and position of the waveguide **508** are configured to ensure that light from each of the eight ray bundles, within the light cone, enters the waveguide **508**. Light from the light cone enters the waveguide **508** via its first planar surface (located nearest the LCOS **502**) and is guided at least partially along the length of the waveguide **508**, before being emitted via its second planar surface, substantially opposite the first surface (located nearest the eye). As will be well understood, the second planar surface is partially reflective, partially transmissive. In other words, when each ray of light travels within the waveguide **508** from the first planar surface and hits the second planar surface, some of the light will be transmitted out of the waveguide **508** and some will be reflected by the second planar surface, back towards the first planar surface. The first planar surface is reflective, such that all light that hits it, from within the waveguide **508**, will be reflected back towards the second planar surface. Therefore, some of the light may simply be refracted between the two planar surfaces of the waveguide **508** before being transmitted, whilst other light may be reflected, and thus may undergo one or more reflections, (or 'bounces') between the planar surfaces of the waveguide **508**, before being transmitted.

[0169] FIG. **5** shows a total of nine "bounce" points, B0 to B8, along the length of the waveguide **508**. Although light relating to all points of the image (V1-V8) as shown in FIG. **4**A is transmitted out of the waveguide at each "bounce" from the second planar surface of the waveguide **508**, only the light from one angular part of the image (e.g. light of one of V1 to V8) has a trajectory that enables it to reach the eye **505**, from each respective "bounce" point, B0 to B8. Moreover, light from a different angular part of the image, V1 to V8, reaches the eye **505** from each respective "bounce" point. Therefore, each angular channel of encoded light reaches the eye only once, from the waveguide **508**, in the example of FIG. **5**.

[0170] The methods and arrangements described above can be implemented in a variety of different applications and viewing systems. For example, they may be implemented in a head-up-display (HUD) or in a head or helmet mounted device (HMD) such as an Augmented Reality (AR) HMD.

[0171] Although virtual images, which require the eye to transform received modulated light in order to form a perceived image, have generally been discussed herein, the methods and arrangements described herein can be applied to real images.

Two-Dimensional Pupil Expansion

[0172] Whilst the arrangement shown in FIG. **5** includes a single waveguide that provides pupil expansion in one dimension, pupil expansion can be provided in more than one dimension, for example in two dimensions. Moreover, whilst the example in FIG. **5** uses a hologram that has been calculated to create channels of light, each corresponding to a different portion of an image, the present disclosure and the systems that are described herebelow are not limited to such a hologram type.

[0173] FIG. **6** shows a perspective view of a system **600** comprising two replicators, **604**, **606** arranged for expanding a light beam **602** in two dimensions.

[0174] In the system **600** of FIG. **6**, the first replicator **604** comprises a first pair of surfaces, stacked parallel to one another, and arranged to provide replication—or, pupil expansion—in a similar manner to the waveguide **508** of FIG. **5**. The first pair of surfaces are similarly (in some cases, identically) sized and shaped to one another and are substantially elongate in one direction. The collimated light beam **602** is directed towards an input on the first replicator **604**. Due to a process of internal reflection between the two surfaces, and partial transmission of light from each of a plurality of output points on one of the surfaces (the upper surface, as shown in FIG. **6**), which will be familiar to the skilled reader, light of the light beam **602** is replicated in a first direction, along the length of the first replicator **604**. Thus, a first plurality of replica light beams **608** is emitted from the first replicator **604**, towards the second replicator **606**.

[0175] The second replicator **606** comprises a second pair of surfaces stacked parallel to one another, arranged to receive each of the collimated light beams of the first plurality of light beams **608** and further arranged to provide replication—or, pupil expansion—by expanding each of those light beams in a second direction, substantially orthogonal to the first direction. The first pair of surfaces are similarly (in some cases, identically) sized and shaped to one another and are substantially rectangular. The rectangular shape is implemented for the second replicator in order for it to have length along the first direction, in order to receive the first plurality of light beams **608**, and to have length along the second, orthogonal direction, in order to provide replication in that second direction. Due to a process of internal reflection between the two surfaces, and partial transmission of light from each of a plurality of output points on one of the surfaces (the upper surface, as shown in FIG. **6**), light of each light beam within the first plurality of light beams **608** is replicated in the second direction. Thus, a second plurality of light beams **610** is emitted from the second replicator **606**, wherein the second plurality of light beams **610** comprises replicas of the input light beam **602** along each of the first direction and the second direction. Thus, the second plurality of light beams **610** may be regarded as comprising a two-dimensional grid, or array, of replica light beams.

[0176] Thus, it can be said that the first and second replicators **604**, **605** of FIG. **6** combine to provide a two-dimensional replicator (or, "two-dimensional pupil expander").

Improved Waveguide

[0177] As described in relation to FIG. **5**, light in the waveguide **508** is reflected between a partially reflective, partially transmissive surface and a reflective surface of a waveguide. Light may undergo one or more reflections or bounces between the two reflective/reflective-transmissive planar surfaces and, at each bounce point on the partially transmissive surface, the light is divided such that a portion of the light is emitted out of the waveguide and the remaining (typically larger) portion of the light is reflected to continue to propagate between the two surfaces of the waveguide. This effectively results in the partially transmissive surface of the waveguide providing a plurality, n, of light emission zones for light waveguided between the first surface and second surface. After each bounce point/emission zone, the intensity of the light propagating in the waveguide will decrease. In other words, the intensity of the light propagating in the waveguide decreases in the direction of waveguiding.

[0178] It is desirable for the intensity of the light emitted out of the waveguide at each of the n light emission zones to be substantially the same. This can be achieved this by providing an improved

waveguide in which a layered coating is provided on the partially transmissive surface of the waveguide to cause the transmissivity of the partially transmissive surface to decrease in the direction of waveguiding. This accounts for the decrease in the intensity of the propagating light in the direction of waveguiding.

[0179] FIG. **7** is schematic cross-sectional view of a waveguide **708** according to the disclosure. The waveguide **708** comprises a first surface **702** and a second surface **704**. A light field **706** (represented by one light ray in FIG. **7**) is shown propagating through the waveguide **708**. The second surface **704** comprises an input port arranged to receive the light field. The first surface **702** is partially transmissive, partially reflective and comprises a coating **703**. The term "coating" is merely used herein for convenience and the person skilled in the art will appreciate that components described as a "coating" may be formed by any method including, but not limited to, a coating process. The second surface **704** is substantially fully reflective (other than at the input). FIG. **7** shows the path of the light field through the waveguide, bouncing between the first and second surfaces. On each reflection at the first surface, the light field divides such that a portion of the light field is emitted through the first surface and a remaining portion is reflected and continues to propagate between the first and second surfaces by reflection. So, an emission zone is effectively formed at each reflection point. FIG. **7** shows 6 emission zones, however the skilled person will understand that there could, of course, be a larger or smaller number of reflections and emission zones. FIG. **7** is merely illustrative.

[0180] The coating **703** comprises a plurality of layers of a first dielectric and a plurality of layers of a second dielectric in an alternating configuration. This is illustrated in FIG. **8**. The layers of the coating will be referred to herein by number, with the layer in contact with the first surface **702** being the first layer (layer **901**). Layer **902** is on top of layer **901** and layer **903** is on top of layer **902**. The layer furthest from the first surface **702**, which is in top of layer **903**, is the fourth layer **904**. In this example, layers **901** and **903** are formed of silicon dioxide (SiO.sub.2) and layers **902** and **904** are formed of titanium dioxide (TiO.sub.2) such that the layers are in an alternating configuration in which subsequent layers of SiO.sub.2 (the first dielectric) are separated by layers of TiO.sub.2 (the second dielectric).

[0181] Each of layers **901** to **904** has a varying thickness in the direction of waveguiding (from left to right in FIG. **9**) and, in this embodiment by way of example, has a linear profile. In other words, the rate of change of thickness of each layer is constant. The profile of each layer can be characterized using a percentage change in thickness. Each layer has a first end **906** and a second end **908**. The percentage change in thickness is defined as the change in thickness from the first end **906** to the second end **908** divided by the thickness at the first end **906** multiplied by 100. For the case of the first layer **901**, the percentage change in thickness is 100×(final thickness **912**−initial thickness **910**)/initial thickness **910**.

[0182] Layer **903** has the same percentage change in thickness as layer **901**. Furthermore, the percentage change value is positive for both layers **901** and **903** (i.e. the thickness of the layer increases from the first end **906** to the second end **908**). Layer **902** and **904** both have different percentage changes to one another and to layers **901** and **903**. Furthermore, both the percentage change of layers **902** and **904** is negative (i.e. the thickness of the layers decreases from the first end **906** to the second end **908**).

[0183] The inventors have found that by, selecting an appropriate number of alternating layers of the first and second dielectric, with the layers having an appropriate thickness and percentage change in thickness from the first end to the second end, a first surface of the waveguide can be provided having a transmissivity that increases in the direction of waveguiding. In this way, the intensity of the light field emitted at each emission zone (i.e. the intensity of each replica emitted at each emission zone) is substantially constant. This may advantageously achieve a substantially spatially homogeneous emission of light from the waveguide.

[0184] An ideal exponential increase of the transmissivity of the first surface **702** is shown in FIG.

9 which is a graph showing transmissivity on the Y axis and position along the first surface **702** on the X axis. The numbers on the X axis represent the n emission zone. Specifically, the transmissivity increases according to the following equation:

[00002]$T(n) = \frac{T(n\text{-}1)}{[1\text{-}T(n\text{-}1)]\times[1\text{-}L]}$

where L is the optical loss factor of the waveguide material.

[0185] The example shown in FIG. **8** is representative only. FIG. **8** is not drawn to scale. Typically, the coatings according to the present invention comprise more than four layers. Herebelow are disclosed **9** examples of layered coatings according to the disclosure which provide a satisfactory increase in the transmissivity of the first surface. Each of these examples have been determined by simulation. The skilled person will appreciate that the present disclosure is not limited to the examples disclosed herein. For example, the coating could comprise a different number of layers (greater or lesser) than the coatings disclosed herein. Furthermore, layers of different dielectric materials could be used.

Example Waveguide

[0186] FIGS. **10**A, **10**B and **10**C relate to a first real example of a coating according to the invention of the present disclosure. The coating of this first example comprises 12 alternating layers of SiO.sub.2 and TiO.sub.2. The layers are numbered from 1 to 12 such that light emitted from the waveguide passes through each of layers 1 to 12 sequentially. The odd numbered layers (layers L1, L3, L5, L7, L9 and L11) are layers of SiO.sub.2. The even numbered layers (layer L2, L4, L6, L8, L10 and L12) are layers of TiO.sub.2.

[0187] FIGS. **10**A and **10**B show how the thickness of each of the layers varies with distance along the first surface of the waveguide (i.e. from the first end to the second end of each layer). The Y axis represents thickness and the X axis represents position in the direction of waveguiding. The numbers 1 to 10 on the X axis represent the emission zones of the first surface. FIG. **10**A shows the SiO.sub.2 layers (i.e. the odd numbered layers in this example). FIG. **10**B shows the TiO.sub.2 layers (i.e. the even numbered layers in the example).

[0188] As shown in FIG. **10**A, there are two subsets of layers of SiO.sub.2. A first subset of the layers has a first value of percentage change of thickness. These layers are represented by the continuous (i.e. unbroken) lines in FIG. **10**A. A second subset of the layers has a second value of percentage change of thickness that is different to the first. These layers are represented by the dashed (i.e. broken) lines in FIG. **10**A. Both the first and second values in this example are negative. In this example, the first value of percentage change is −50%+/−10%. The second value of percentage change is −60%+/−10%. For the avoidance of doubt, the first value is different to the second value. The minimum thickness of the thinnest layer is 50 nm +/−15 nm.

[0189] As shown in FIG. **10**B, there are two subsets of layers of TiO.sub.2. A first subset of the layers has a third value of percentage change of thickness. These layers are represented by the continuous lines in FIG. **10**B. A second subset of the layers has a fourth value of percentage change of thickness that is different to the first. These layers are represented by the dashed lines in FIG. **10**B. Both the first and second values in this example are positive. In this example, the third value of percentage change is 20%+/−10%. The fourth value of percentage change is 60%+/−10%. The minimum thickness of the thinnest layer is 50 nm+/−15 nm.

[0190] The skilled person will appreciate that, while the percentage change in the thickness of two layers may be the same, the maximum and minimum thicknesses of those layers may differ. For example, a first layer having a minimum thickness of 2 nm and a maximum thickness of 4 nm has a percentage increase of 100%. A second layer having a minimum thickness of 5 nm and a maximum thickness of 10 nm also has a percentage increase of 100% despite the absolute thicknesses of the two layers being different.

[0191] Coated waveguides according to the present disclosure will typically perform differently at different wavelengths. As the skilled person will appreciate, a combination of blue, green and red can be used to provide full colour images (i.e. light having three different wavelengths). For

example, in many holographic systems, blue, green and red light sources are used to generate the full colour image. If the waveguide of the present disclosure is to be suitable for use in such systems, the transmissivity of the coated first surface should substantially follow the ideal transmissivity shown in FIG. **9** at each desired wavelength.

[0192] FIG. **10**C is a graph showing the transmissivity of the first surface of the waveguide of the First Example. The Y axis represents transmissivity. The X axis represents the nth emission zone along the first surface of the waveguide. The solid (unbroken) line **1002** of FIG. **10**C represents the ideal transmissive behaviour (also shown in FIG. **9**). The three dashed (broken) lines show the simulated transmissivity of the waveguide at three different wavelengths of electromagnetic radiation. The three wavelengths shown in FIG. **10**C relate to blue **1004**, green **1006** and red **1008** laser light, respectively. FIG. **10**C shows how the coating of the first example provides good transmissive performance at each of the three wavelengths.

[0193] The inventors have appreciated that, when confining the problem of providing a changing transmissivity behavior to only three specific wavelengths, an alternating stack of layers of the first and second dielectric can provide a desired increase in transmissivity in the waveguiding direction even when the number of discrete values of percentage change in thickness is less than (preferably significantly less than) the total number of layers. This is advantageous because it may be simpler, cheaper and more reliable to manufacture such a coating compared to a coating comprising a unique percentage change of thickness for each layer. This is explained with respect to the First Example below.

[0194] The First Example comprises 12 layers of dielectric. However, the percentage change in the thickness of each layer from the first end to the second end of that layer has one of four discrete allowable values. In particular, each layer of the first dielectric has a percentage change in thickness equal to either a first value or a second value and each layer of the second dielectric has a percentage change in thickness equal to either a third value or a fourth value. So, the total number of layers of the first and second dielectric (i.e. 12) is greater than the total number of discrete allowable values (i.e. four). As demonstrated in FIG. **10**C, the transmissivity at blue, green and red wavelengths of that the coating is designed for is acceptable. However, as shown in FIG. **11**, the transmissivity of the coated waveguided does not perform acceptably at other wavelengths.

[0195] FIG. **11** is a graph. The Y axis of FIG. **11** represents transmissivity. The X axis represents the nth emission zone along the first surface of the waveguide. The solid line **1102** represents the ideal transmissivity of the first surface. The solid line **1102** appears to differ to that shown in FIG. **9** but this is because the scale of the Y axis in FIG. **11** is different to that of FIG. **9**. The three dashed lines show the simulated transmissivity of the waveguide at three different wavelengths of electromagnetic radiation that the waveguide of the First Example has not been configured for (**1104**, **1106**, **1108**). The three dashed lines show that the transmissivity of the coated waveguide does not follow the ideal behaviour at the respective wavelengths. The **1104** wavelength diverges very significantly from the ideal behaviour.

[0196] So, the transmissivity of the waveguide having the coating of First Example for wavelengths outside of the specific wavelengths that the waveguide coating is designed for may not be acceptable. But that is not a concern for applications in which those wavelengths are not transmitted through the waveguide (i.e. are not used for formation of the image).

[0197] Further real examples of a coating can be found in in GB patent application 2214069.3 (published as GB2623054A), which is hereby incorporated in its entirety by reference.

Method of Manufacture

[0198] An advantage of the layered coating of the present disclosure is that it can be manufactured in an inexpensive, fast and reliable way. Herein, one example of such a method is disclosed. However, the skilled person will appreciate that other methods are possible.

[0199] The method comprises the steps of providing a waveguide substrate comprising a pair of parallel surfaces arranged to provide waveguiding therebetween; and applying a plurality of layers

of a first dielectric and a plurality of layers of a second dielectric to the waveguide substrate such that the layers of first and second dielectric are in an alternating configuration.

[0200] An apparatus for performing the method comprises two sources of a first dielectric material (for example SiO.sub.2) and two sources of a second dielectric material (for example TiO.sub.2), a shadow mask comprising first to fourth trapezoidal apertures and a means for moving a waveguide substrate with respect to the shadow mask. Each aperture of the shadow mask is associated with a source of dielectric material. During manufacture of each layer, dielectric material from the one of the sources is configured to flow through one of the apertures of the shadow mask. A waveguide substrate is moved with respect to the shadow mask (or vice versa) such that the shadow mask is between the source and the waveguide substrate. The waveguide substrate passes under the aperture and a layer of dielectric material is formed on a surface of the waveguide substrate. This is shown in FIGS. **19**A, **19**B and **19**C.

[0201] FIG. **12** is a schematic cross-sectional view of a portion of the apparatus for manufacturing a waveguide according to the present disclosure, the cross-section being in X-Y plane. The portion of the apparatus shown in FIG. **12** comprises a first source of SiO.sub.2 and a shadow mask **2004** comprising a first aperture **2006**. A waveguide substrate **2008** (in the form of a glass or Perspex block or slab) is also shown. SiO.sub.2 material **2010** is configured to flow out of the first source **2002** and through the first aperture **2006** of the shadow mask **2004**. The flow of material is in the negative Y direction. The shape of the first aperture **2006** determines the shape of the flow SiO.sub.2 downstream of the shadow mask **2004**. FIG. **12** is not drawn to scale.

[0202] The means to move the waveguide substrate **2008** (not shown in drawings) is arranged to move the waveguide substrate **2008** in a first plane that is perpendicular to the Y direction such that the waveguide substrate **2008** passes under the first aperture **2006**. In some embodiments, the motion is entirely in the X direction. However, in other embodiments the waveguide substrate **2008** may be rotated in the first plane such that the motion is in both an X direction and a Z direction.

[0203] In FIG. **12**, the waveguide substrate **2008** has yet to pass under the first aperture **2006** and so no dielectric coating layer is present on the waveguide substrate **2008**. As the waveguide substrate passes under the first aperture **2006** (in the X direction), dielectric coating is deposited on the substrate. Generally, in order to manufacture a complete layer, the waveguide substrate **2008** will need to pass under the first aperture **2006** multiple times until the desired thickness is achieved.

[0204] Once the first layer **2010** has been formed, the waveguide substrate **2008** will be moved under one of the sources of second dielectric material such that a second layer of (of second dielectric material) is formed on top of the first layer **2010** of first dielectric material.

[0205] FIG. **13** is a schematic view of the shadow mask **2004**, the view being taken in the X-Z plane (i.e. orthogonal to the plane of view of FIG. **12**). The shadow mask comprises four apertures, the first aperture **2006** (described above), a second aperture **2102**, a third aperture **2104** and a fourth aperture **2106**. The first and third apertures **2006**, **2104** are each connected to source(s) of the first dielectric material. The second and fourth aperture **2102** and **2106** are each connected to source(s) of the second dielectric material. Each of the apertures has a trapezoidal shape comprising a short base and a long base.

[0206] In order to manufacture subsequent alternating layers of the first and second dielectric material, the waveguide substrate **2008** is moved under different apertures in turn. The order of the layers can be controlled depending on the order of the apertures that the waveguide substrate **2008** is moved under. The waveguide substrate **2008** is rotated/moved with respect to the shadow mask **2004** in order to pass under different apertures such that the waveguide substrate **2008** is passed under the aperture in the X direction and such that short base and long base of the aperture spaced apart in the Z direction. In this way, the thickness of each deposited layer in the Z direction varies. A percentage change in the thickness of the layer from a first end to a second end (in the Z direction) will depend on the percentage change in width of the short base and the long base of the

respective aperture. This is explained with reference to FIG. **13**.

[0207] FIGS. **14**A, **14**B, **14**C and **14**D are each cross-sectional schematic views of four different first layers formed on a waveguide substrate **2008**. FIG. **14**A shows a layer **2202** formed when the waveguide substrate **2008** is passed under first aperture **2006**. FIG. **14**B shows a layer **2204** formed when the waveguide substrate **2008** is passed under second aperture **2102**. FIG. **14**C shows a layer **2206** formed when the waveguide substrate **2008** is passed under second aperture **2104**. FIG. **14**D shows a layer **2208** formed when the waveguide substrate **12008** is passed under second aperture **2106**. As the skilled person will appreciate, the thickness profile of each of layers **2202** to **2208** corresponds to the shape of the respective aperture.

[0208] In particular, layers **2202** and **2206** have a positive gradient from left to right because the width of the apertures **2006** and **2104** increase in the Z direction of the waveguide substrate when the shadow mask and the waveguide substrate are rotated with respect to one another. The percentage change in thickness of the layer **2206** is greater than that of layer **2202** because the percentage change of the width of the short base to the long base of aperture **2104** is greater than aperture **2006**. Layers **2204** and **2208** have a negative gradient from left to right because width of apertures **2102** and **2106** decrease in the Z direction when the shadow mask and the waveguide substrate are rotated with respect to one another. The percentage change in thickness of the layer **2204** is greater than that of layer **2208** because the percentage change of the width of the short base to the long base of aperture **2102** is greater than aperture **2106**.

[0209] As will be appreciated, the arrangement of the shadow mask **2004** with four differently shaped apertures and with two apertures connected to a first dielectric material source and two apertures connected to a second dielectric material source provides a means for manufacturing a plurality of layers of a first dielectric and a plurality of layers of second dielectric arranged in an alternating configuration. Each layer of the first dielectric will have a percentage change in thickness equal to either a first value or a second value, wherein all layers having the first value are related to, for example, the first aperture **2006** and all layers having the second value are related to, for example, the second aperture **2102**. Each layer of the second dielectric will have a percentage change in thickness equal to either a first value or a second value, wherein all layers having the third value are related to, for example, the third aperture **2104** and all layers having the second value are related to, for example, the fourth aperture **2106**. In other words, a means for manufacturing a layered coating according to the present disclosure is provided.

[0210] The advantages of this method of manufacture are that the waveguide substrate can quickly and simply be moved/rotated to pass under the apertures of the shadow mask **2004** as necessary to build up a plurality of alternating layers of the first and second dielectric material, as necessary. The number of unique masks having a different shape governs the number of discrete values for the percentage change in thickness of layers that are available and so the percentage change in thickness of any particular layer can be controlled simply by selecting the aperture order. The absolute thickness of any layer can be controlled by either controlling the speed at which the waveguide substrate passes under a particular shadow mask **2004** or be controlling the material flow speed.

[0211] It should be understood that the manufacturing method disclosed herein is not restricted to their being four apertures and two sources of dielectric material. For example increasing or decreasing the number of differently shaped masks will simply have the effect of increasing or decreasing the discrete number of allowable percentage change in thickness values of layers available.

Improved Method of Coating Application

[0212] In overview, there is provided a further improved method of coating design includes the following steps: [0213] 1. Designing the base structure, which is usually a multilayer mirror with RGB windows to be used at the lowest T % position. There is no gradient involved in this step.

[0214] 2. Apply linear gradient on the multilayer base structure, optimising the gradient so that

coating transmission at different positions can meet any requirements.

[0215] The general idea is to use the gaps among RGB windows as design freedom to help maintaining a stable and synchronised increase of transmission within RGB windows. As a result, the design is very sensitive to gradient—a small amount of change (several percentage) may change coating transmission drastically.

[0216] In practice, the inventor found it was very difficult to adjust the mask to meet the same gradient thickness as designed. There are several challenges: [0217] Adjusting the mask is a time-consuming process, which requires precise machining of metal. [0218] There is no simple relationship between gradient thickness and the physical gradient of the mask opening. To achieve linear gradient thickness, the mask usually needs to have some non-linearity. It is typically a trial-and-error process to get close to the required gradient after several iterations. [0219] In mass production, masks will degrade in weeks and need maintenance/re-conditioning. This increases the chances of small deviations in gradient thickness.

[0220] What usually happens is that the gradient masks are close to the design but not exactly accurate. To help work around these imperfect gradient masks, there is disclosed herein a 'post-process' coating optimisation, which does not need any hardware change but can deliver significant improvement to the coating performance. The method comprises: [0221] Fixing the gradient ("coating function") as is—that is, accepting the small deviation, with no further optimisation. [0222] Finding the best sets of multilayer thicknesses that can 'fully utilise' the fixed gradient ("coating function").

[0223] Prior to the optimisation of the present disclosure, the multilayer thicknesses were designed for the ideal gradient, thus cannot fit the actual gradient. This optimisation finds a (slightly) different set of thicknesses (one example of a "coating parameter") to work with the actual gradient (measured "coating function").

[0224] FIG. **15** shows an improved method of applying the coating (for example the coatings described above) to the waveguide substate **2008** using the apparatus shown in FIGS. **12**, **13**, **14**A, **14**B, **14**C and **14**D and described above.

[0225] First, an optimised transmissivity at a plurality of locations along the length of the waveguide (for example at "bounce" points B0 to B8 in FIG. **5** or the emission zones of FIGS. **7**, **10**A, **10**B, **10**C and **11**) is calculated. From this, a determination **3000** of the number of layers of coating and the profile of each layer to achieve this optimised transmissivity is carried out (as discussed above), along with the determination of the particular shadow mask **2004**, waveguide substrate **2008** speed and material flow speed required to apply said optimised layer formation (also as discussed above). This determination **3000** outputs optimised values **3100**.

[0226] Using these optimised values **3100**, the layers of coating are then applied to the waveguide substate **2008** as described above in a formation step **3200**. This forms a first coated waveguide. Given factors such as the wear on the shadow mask **2004**, it has been observed that the actual transmissivity of the first waveguide will be different to that optimised during the determination **3000**. This is confirmed by a measurement **3300** of the thicknesses of the layers at the plurality of locations used in the determination **3000** (a change in thickness of any layer at any point along the waveguide resulting in a change of transmissivity, as discussed above). As explained above, the step of measurement **3300** may use data from a pass-or-fail assessment of the coating. The present disclosure, however, relates to using the measured data to determine an actual coating function of the mask rather than simply assessing pass-or-fail.

[0227] Using the measurement **3300** values, the optimisation completed for the initial determination **3000** is repeated to account for the difference between the simulated optimised waveguide and the one actually produced in the formation step **3200**. As such, the values produced by this re-optimisation can be used to form layers of coating that take into account wear or other imperfections on the shadow mask **2004**, maintaining the required optical properties. Notably, the re-optimisation leverages a degree of freedom associated with the coating parameter—e.g. coating

speed or minimum thickness—to provide better performance without replacing the mask.

[0228] In some embodiments of the method, an optimisation loop can be formed, as shown in FIG. **16**. In this case, repeated optimisation **3400** produces a new set of values **3500** relating to an updated waveguide substrate **2008** speed and material flow speed required to apply said updated layer formation. The intended profile of each layer is kept the same to avoid having to replace the worn shadow mask **2004** with one having differently shaped and/or unworn apertures.

[0229] Then, a measurement **3600** is taken of the actual transmissivity of the first coated waveguide across the points along the waveguide. This is compared to the optimised transmissivity of the optimised values **3100**, in a comparison step **3700**. If this actual transmissivity is within a predetermined tolerance of the optimised values **3100** (i.e., within an acceptable error rate), then mass production **3800** of the waveguides begins or continues (until the shadow mask **2004** wears further and the process needs to be repeated). However, if the actual transmissivity is not within this tolerance a second coated waveguide is formed in formation step **3200** using the new set of values **3500**. The loop of steps **3200**, **3300**, **3400** and **3600** is repeated until said tolerance is met. As such, a set of waveguides are produced with the required transmissivity properties regardless of wear of the shadow mask **2004** introducing inaccuracies in the application of the coating. For the avoidance of doubt, the present disclosure is characterised by deriving an "actual" coating function of each mask by measurement (e.g. using the pass-or-fail test data) and using a degree of freedom associated with the coating parameter to optimise performance and extend the lifetime of the mask. The effectiveness of this approach is described below in relation to FIGS. **18**A, **18**B, **18**C, **19**A, **19**B and **19**C.

[0230] FIG. **17** shows an alternate method of applying the coating. First, as with the method described above, an optimised transmissivity at a plurality of locations along the length of the waveguide (for example at "bounce" points B0 to B8 in FIG. **5** or the emission zones of FIGS. **7** and **10**A to **11**) is calculated. From this, a determination **4000** of the number of layers of coating and the profile of each layer to achieve this optimised transmissivity is carried out (as discussed above), along with the determination of the particular shadow mask **2004**, waveguide substrate **2008** speed and material flow speed required to apply said optimised layer formation (also as discussed above). This determination **4000** outputs optimised values **4100**.

[0231] Using these optimised values **4100**, the first layer of coating is applied to the waveguide substate **2008** as described above in a formation step **4200**. As discussed in relation to the method above, given factors such as the wear on the shadow mask **2004**, it is expected that the actual transmissivity of the first layer will be different to that optimised during the determination **4000**. This is confirmed by a measurement **4300** of the thicknesses of the first layer at the plurality of locations used in the determination **4000** (a change in thickness resulting in a change of transmissivity, as discussed above).

[0232] Using the measurement **4300** values, a further optimisation **4400** is completed to produce a new set of values **4500** relating to an updated waveguide substrate **2008** speed and material flow speed required to apply the second layer whilst taking into account the difference between the simulated optimised first layer and the one actually produced in the formation step **4200**. As with the method above, the intended profile of each layer is kept the same to avoid having to replace the worn shadow mask **2004** with one having differently shaped and/or unworn apertures.

[0233] Finally, in a decision step **4600**, if this is the last layer to be applied, then the process finishes and the coating has been completed. However, if this is not the last layer, then the next layer is formed in formation step **4200** using the new set of values **4500**. The loop of steps **4200**, **4300** and **4400** is repeated until the coating is completed. As such, a waveguide is produced with the required transmissivity properties regardless of wear of the shadow mask **2004** introducing inaccuracies in the application of layers of the coating.

Examples of Coatings Applied Using the Improved Method

[0234] In overview, FIGS. **18**A, **18**B and **18**C show the simulation of original multilayer

thicknesses with some actual, deviated gradient thicknesses. FIG. **18**C shows that half of the gradient thicknesses (the even numbered layers) are no longer linear, and the actual gradient amplitude is also higher than the designed. The other half thicknesses (FIG. **18**A-odd numbered layers) remain linear in this example. As a result, shown in FIG. **18**C, the RGB transmission deviates from the target line **5002**. In overview, in FIGS. **19**A, **19**B and **19**C, all gradient percentage (masks/"coating functions") remain the same, but layer thicknesses ("coating parameters) are re-optimised—i.e. curves moving up/down slightly in bottom left/right plots but maintaining the same shape. After optimisation, the graded RGB transmission is much closer to the original target. Once the optimisation algorithm is finalised, the optimisation can run automatically with any gradient input (within a certain range) and returns optimal thicknesses in near real time.

[0235] More specifically, FIGS. **18**A, **18**B and **18**C relate to a real example of a coating applied using a known method by a shadow mask **2004** that has worn causing inaccuracies. The coating of this example comprises 30 alternating layers of two different dielectric materials. The layers are numbered from 1 to 30 such that light emitted from the waveguide passes through each of layers 1 to 30 sequentially.

[0236] FIGS. **18**A and **18**B show how the actual, measured thickness of each of the layers varies with distance along the first surface of the waveguide (i.e. from the first end to the second end of each layer). In accordance with this disclosure, these actual thicknesses are different to what was specified. That is, they are different to the designed or required layers. This deviation may be due to aging of the mask. The Y axis represents thickness and the X axis represents position in the direction of waveguiding. The numbers 1 to 9 on the X axis represent the emission zones of the first surface.

[0237] Coated waveguides according to the present disclosure will typically perform differently at different wavelengths. As the skilled person will appreciate, a combination of blue, green and red can be used to provide full colour images (i.e. light having three different wavelengths). For example, in many holographic systems, blue, green and red light sources are used to generate the full colour image. If the waveguide created by the present disclosure is to be suitable for use in such systems, the transmissivity of the coated first surface should substantially follow the ideal transmissivity shown in FIG. **9** at each desired wavelength.

[0238] FIG. **18**C is a graph showing the transmissivity of the first surface of the waveguide of this example. FIG. **18**C is obtainable by optical measurement of the transmissivity of layers L1 to L30 or by simulation. The Y axis represents transmissivity. The X axis represents the nth emission zone along the first surface of the waveguide. The solid (unbroken) line **5002** of FIG. **18**C represents the ideal transmissive behaviour (also shown in FIG. **9**). The three dashed (broken) lines show the simulated transmissivity of the waveguide at three different wavelengths of electromagnetic radiation. The three wavelengths shown in FIG. **18**C relate to blue **5004**, green **5006** and red **5008** laser light, respectively. It can be seen that the three wavelengths **5004**, **5006**, **5008** differ substantially from the ideal transmissible behaviour **5002** in several of the emission zones. This is caused by the imperfect performance (e.g. wear) on the shadow mask **2004** as discussed above.

[0239] FIGS. **19**A, **19**B and **19**C relate to a real example of the coating of FIGS. **18**A, **18**B and **18**C applied according to the method of the present disclosure—i.e. that has been further optimised in accordance with the present disclosure. The coating of this example comprises 30 alternating layers of two different dielectric materials. The layers are numbered from 1 to 30 such that light emitted from the waveguide passes through each of layers 1 to 30 sequentially.

[0240] FIGS. **19**A and **19**B show how the actual, measured thickness of each of the layers varies with distance along the first surface of the waveguide (i.e. from the first end to the second end of each layer) after re-optimisation (or the step of second optimisation) in accordance with this disclosure. The Y axis represents thickness and the X axis represents position in the direction of waveguiding. The numbers 1 to 9 on the X axis represent the emission zones of the first surface. As can be seen, the thickness of each layer L1 to L30 has been shifted relative to the corresponding

thicknesses shown in FIGS. **18**A and **18**B.

[0241] For the avoidance of doubt, the layers L1 to L30 represented by FIGS. **19**A and **19**B have been determined by using at least one measured coating function after a first coating run and allowing the coating parameter to float during design in order to improve the achieved transmissivity shown in FIG. **19**C.

[0242] FIG. **19**C is a graph showing the transmissivity of the first surface of the waveguide of this example. The Y axis represents transmissivity. The X axis represents the nth emission zone along the first surface of the waveguide. The solid (unbroken) line **6002** of FIG. **19**C represents the ideal transmissive behaviour (also shown in FIG. **9**). The three dashed (broken) lines show the simulated transmissivity of the waveguide at three different wavelengths of electromagnetic radiation. The three wavelengths shown in FIG. **19**C relate to blue **6004**, green **6006** and red **6008** laser light, respectively. As can be seen, the adjusted thicknesses of each layer L1 to L30 have caused the wavelengths **6004**, **6006**, **6008** to be closer to the ideal transmissible behaviour **6002** across the length of the waveguide than that produced by the known method (as seen in FIG. **18**C). This improved transmissivity has been produced without replacing or repairing worn shadow mask **2004**.

Additional Features

[0243] Embodiments refer to an electrically-activated LCOS spatial light modulator by way of example only. The teachings of the present disclosure may equally be implemented on any spatial light modulator capable of displaying a computer-generated hologram in accordance with the present disclosure such as any electrically-activated SLMs, optically-activated SLM, digital micromirror device or microelectromechanical device, for example.

[0244] In some embodiments, the light source is a laser such as a laser diode. In some embodiments, the detector is a photodetector such as a photodiode. In some embodiments, the light receiving surface is a diffuser surface or screen such as a diffuser. The holographic projection system of the present disclosure may be used to provide an improved head-up display (HUD). In some embodiments, there is provided a vehicle comprising the display system installed in the vehicle to provide a HUD. The vehicle may be an automotive vehicle such as a car, truck, van, lorry, motorcycle, train, airplane, boat, or ship.

[0245] The quality of a holographic reconstruction may be affected by the so-called zero order problem which is a consequence of the diffractive nature of using a pixelated spatial light modulator. Such zero-order light can be regarded as "noise" and includes for example specularly reflected light, and other unwanted light from the SLM.

[0246] In the example of Fourier holography, this "noise" is focussed at the focal point of the Fourier lens leading to a bright spot at the centre of the holographic reconstruction. The zero order light may be simply blocked out however this would mean replacing the bright spot with a dark spot. Some embodiments include an angularly selective filter to remove only the collimated rays of the zero order. Embodiments also include the method of managing the zero-order described in European patent 2,030,072, which is hereby incorporated in its entirety by reference.

[0247] In some embodiments, the size (number of pixels in each direction) of the hologram is equal to the size of the spatial light modulator so that the hologram fills the spatial light modulator. That is, the hologram uses all the pixels of the spatial light modulator. In other embodiments, the hologram is smaller than the spatial light modulator. More specifically, the number of hologram pixels is less than the number of light-modulating pixels available on the spatial light modulator. In some of these other embodiments, part of the hologram (that is, a continuous subset of the pixels of the hologram) is repeated in the unused pixels. This technique may be referred to as "tiling" wherein the surface area of the spatial light modulator is divided up into a number of "tiles", each of which represents at least a subset of the hologram. Each tile is therefore of a smaller size than the spatial light modulator. In some embodiments, the technique of "tiling" is implemented to increase image quality. Specifically, some embodiments implement the technique of tiling to

minimise the size of the image pixels whilst maximising the amount of signal content going into the holographic reconstruction. In some embodiments, the holographic pattern written to the spatial light modulator comprises at least one whole tile (that is, the complete hologram) and at least one fraction of a tile (that is, a continuous subset of pixels of the hologram).

[0248] In embodiments, only the primary replay field is utilised and system comprises physical blocks, such as baffles, arranged to restrict the propagation of the higher order replay fields through the system.

[0249] In embodiments, the holographic reconstruction is colour. In some embodiments, an approach known as spatially-separated colours, "SSC", is used to provide colour holographic reconstruction. In other embodiments, an approach known as frame sequential colour, "FSC", is used.

[0250] The method of SSC uses three spatially-separated arrays of light-modulating pixels for the three single-colour holograms. An advantage of the SSC method is that the image can be very bright because all three holographic reconstructions may be formed at the same time. However, if due to space limitations, the three spatially-separated arrays of light-modulating pixels are provided on a common SLM, the quality of each single-colour image is sub-optimal because only a subset of the available light-modulating pixels is used for each colour. Accordingly, a relatively low-resolution colour image is provided.

[0251] The method of FSC can use all pixels of a common spatial light modulator to display the three single-colour holograms in sequence. The single-colour reconstructions are cycled (e.g. red, green, blue, red, green, blue, etc.) fast enough such that a human viewer perceives a polychromatic image from integration of the three single-colour images. An advantage of FSC is that the whole SLM is used for each colour. This means that the quality of the three colour images produced is optimal because all pixels of the SLM are used for each of the colour images. However, a disadvantage of the FSC method is that the brightness of the composite colour image is lower than with the SSC method—by a factor of about 3-because each single-colour illumination event can only occur for one third of the frame time. This drawback could potentially be addressed by overdriving the lasers, or by using more powerful lasers, but this requires more power resulting in higher costs and an increase in the size of the system.

[0252] Examples describe illuminating the SLM with visible light but the skilled person will understand that the light sources and SLM may equally be used to direct infrared or ultraviolet light, for example, as disclosed herein. For example, the skilled person will be aware of techniques for converting infrared and ultraviolet light into visible light for the purpose of providing the information to a user. For example, the present disclosure extends to using phosphors and/or quantum dot technology for this purpose.

[0253] Some arrangements describe 2D holographic reconstructions by way of example only. In other arrangements, the holographic reconstruction is a 3D holographic reconstruction. That is, in some arrangements, each computer-generated hologram forms a 3D holographic reconstruction.

[0254] The methods and processes described herein may be embodied on a computer-readable medium. The term "computer-readable medium" includes a medium arranged to store data temporarily or permanently such as random-access memory (RAM), read-only memory (ROM), buffer memory, flash memory, and cache memory. The term "computer-readable medium" shall also be taken to include any medium, or combination of multiple media, that is capable of storing instructions for execution by a machine such that the instructions, when executed by one or more processors, cause the machine to perform any one or more of the methodologies described herein, in whole or in part.

[0255] The term "computer-readable medium" also encompasses cloud-based storage systems. The term "computer-readable medium" includes, but is not limited to, one or more tangible and non-transitory data repositories (e.g., data volumes) in the example form of a solid-state memory chip, an optical disc, a magnetic disc, or any suitable combination thereof. In some example

embodiments, the instructions for execution may be communicated by a carrier medium. Examples of such a carrier medium include a transient medium (e.g., a propagating signal that communicates instructions).

[0256] It will be apparent to those skilled in the art that various modifications and variations can be made without departing from the scope of the appended claims. The present disclosure covers all modifications and variations within the scope of the appended claims and their equivalents.

Claims

1. A method of forming a transmission coating for a waveguide, wherein the transmission coating comprises a plurality of layers and the method comprises the steps of: a. determining a first coating parameter and a coating function for each layer to optimise the transmissivity at a plurality of locations along the waveguide for a plurality of different wavelengths, the coating function being chosen from a plurality of allowable coating functions; b. forming the plurality of layers using the determined coating parameters and coating functions; c. measuring a thickness of at least one layer at each of the plurality of locations, wherein the measurements indicate that the coating function deviates from that selected during the optimisation of step a; and d. determining a second coating parameter for at least one layer by repeating the optimisation of step a using the coating function derived from the measurements of step c.

2. A method as claimed in claim 1, wherein each coating function is a percentage change in thickness of the layer from a first end of the waveguide to a second end of the waveguide along the length of the waveguide.

3. A method as claimed in claim 2, wherein the number of allowable coating functions is between two and six, optionally between two and four, optionally four.

4. A method as claimed in claim 2, wherein each determined coating function is linear, optionally wherein the measurements of step c indicate that the coating function is non-linear.

5. A method as claimed in claim 1, wherein the first coating parameter relates to a thickness of the layer at its thinnest point.

6. A method as claimed in claim 1, wherein the formation of the plurality of layers of step b comprises controlling a flow of a coating material towards the waveguide using a shadow mask.

7. A method as claimed in claim 6, wherein the first and second coating parameters are one of: the rate of the flow of coating material towards the waveguide; or a rate at which the waveguide moves through the flow of coating material.

8. A method as claimed in claim 7, wherein step a further comprises determining a third coating parameter and step d further comprises determining a fourth coating parameter, the third and fourth coating parameters being the other of: the rate of the flow of coating material towards the waveguide; or the rate at which the waveguide moves through the flow of coating material.

9. A method as claimed in claim 1, wherein the plurality of layers comprises at least one dielectric.

10. A method as claimed in claim 1, wherein the formation of the plurality of layers of step b comprises forming alternate layers of a first material and a second material; optionally wherein the first material is a first dielectric and the second material is a second dielectric, the first dielectric being a first oxide, fluoride, sulfide or nitrate of a first transition metal or semiconductor and the second dielectric being a second oxide, fluoride, sulfide or nitrate of a second transition metal or semiconductor and/or wherein a difference in refractive index between the first material and second material is greater than 0.4.

11. A method as claimed in claim 1, wherein the plurality of different wavelengths comprises a first wavelength, a second wavelength and a third wavelength, optionally wherein the first wavelength is in the range 630-670 nm, the second wavelength is in the range 500-540 nm and the third wavelength is in the range 430-470 nm.

12. A method as claimed in claim 1, wherein the optimised transmissivity at each of the plurality of

locations, T(n), along the waveguide satisfies the following equation: $T(n) = \frac{T(n-1)}{[1-T(n-1)] \times [1-L]}$
wherein L is an optical loss factor of the waveguide material.

13. A method as claimed in claim 1, wherein the optimisations of steps a and d further comprise accounting for the transmissivity of a further transmission coating on an opposing surface to the transmission coating being applied.

14. A method as claimed in claim 1, wherein the deviated coating function is determined by fitting the original coating function to the measured thickness.

15. A method as claimed in claim 1, wherein the measurement of step c comprises measuring the thickness of each layer.

16. A method as claimed in claim 1, wherein the repeated optimisation of step d comprises adjusting a coating parameter of each layer at each of the plurality of locations.

17. A method as claimed in claim 1, further comprising the steps of: e. forming or simulating the plurality of layers using the deviated coating function and the adjusted coating parameters of step d; f. determining or simulating the transmissivity at the plurality of locations along the waveguide for the plurality of different wavelengths; and g. repeating steps c to f until the determined/simulated transmissivity of step f is within a tolerance of the optimised transmissivity of step a.

18. A transmission coating for a waveguide formed using the method as claimed in claim 1.

19. A waveguide comprising a transmission coating formed using the method as claimed in claim 1 optionally further comprising a reflection coating on an opposing surface to the transmission coating.

20. A holographic system comprising: a display device arranged to display a hologram of an image and to output spatially modulated light in accordance with the hologram; and a first waveguide as claimed in claim **19**; optionally further comprising a second waveguide.