

# INVIS Integrated Night Vision Surveillance and Observation System

Alexander Toet\*, Maarten A. Hogervorst, Judith Dijk, Rob van Son  
TNO Defense, Security and Safety, the Netherlands

## ABSTRACT

We present the design and first field trial results of the all-day all-weather INVIS Integrated Night Vision surveillance and observation System. The INVIS augments a dynamic three-band false-color nightvision image with synthetic 3D imagery in a real-time display. The night vision sensor suite consists of three cameras, respectively sensitive in the visual (400-700 nm), the near-infrared (700-1000 nm) and the longwave infrared (8-14  $\mu\text{m}$ ) bands of the electromagnetic spectrum. The optical axes of the three cameras are aligned. Image quality of the fused sensor signals can be enhanced in real-time through Dynamic Noise Reduction, Superresolution, and Local Adaptive Contrast Enhancement. The quality of the longwave infrared image can be enhanced through Scene-Based Non-Uniformity Correction (SBNUC), intelligent clustering and thresholding. The visual and near-infrared signals are used to represent the resulting multiband nightvision image in realistic daytime colors, using the Color-the-Night color remapping principle. Color remapping can also be deployed to enhance the visibility of thermal targets that are camouflaged in the visual and near-infrared range of the spectrum. The dynamic false-color nighttime images can be augmented with corresponding synthetic 3D scene views, generated in real-time using a geometric 3D scene model in combination with position and orientation information supplied by the GPS and inertial sensors of the INVIS system.

**Keywords:** augmented reality, image fusion, false color, natural color mapping, real-time fusion, night vision

## 1. INTRODUCTION

We present the design and first field trial results of the all-day all-weather INVIS Integrated Night Vision surveillance and observation System. The INVIS integrates the TRICLOBS<sup>12</sup> TRI-band Color Low-light OBServation system with a synthetic 3D scene generation system, advanced image fusion and image processing capabilities. The sensor suite of the system is sensitive in the visual (400-700 nm), the near-infrared (700-1000 nm) and the longwave infrared (8-14  $\mu\text{m}$ ) bands of the electromagnetic spectrum, and comprises two digital image intensifiers and a thermal (LWIR) camera. Position information is provided by a 3D digital GPS system. The viewing direction of the system (the azimuth and tilt of its optical axis) is provided by a digital compass.

Image quality the fused sensor signals can be enhanced through Dynamic Superresolution (DSR) and Local Adaptive Contrast Enhancement (LACE). The quality of the thermal image can be further enhanced through Scene-Based Non-Uniformity Correction (SBNUC), intelligent clustering and thresholding. Either synthetic scene views, or the combination of the two intensified images, representing the visual and near-infrared bands, can be used to display the resulting multiband nightvision image in realistic daytime colors, using the Color-the-Night color mapping principle. This mapping can also be deployed to enhance the visibility of thermal targets that are camouflaged in the visual and near-infrared bands.

For applications such as military command and control, facility security, and catastrophe management, optimal situational awareness is crucial. The INVIS system augments the dynamic false-color nighttime images with corresponding synthetic 3D scene views, generated in real-time using a geometric 3D scene model of the operational area, in combination with the 6 degree-of-freedom position and orientation information supplied by the GPS (Global Positioning System) and inertial sensors of the system. The INVIS is designed to serve as an all-day all-weather navigation and surveillance tool.

---

\*lex.toet@tno.nl; phone +31-346-356237; fax +31-346-353977; <http://lextoet.googlepages.com/>

## 2. SYSTEM HARDWARE

The INVIS integrates the TRICLOBS<sup>12</sup> TRI-band Color Low-light OBServation system with a synthetic image generation system, and advanced image fusion and processing capabilities. The INVIS system comprises a three-band nightvision sensor suite, GPS and inertial sensors, a geometric 3D scene model, and a laptop to dynamically process (i.e. fuse, colorize, and enhance) the imagery and to generate synthetic images from the 3D scene model. The whole system is portable and can either run on an internal battery pack, or on 220V AC. The next sections describe each of the system hardware components.

### 2.1. Sensor Suite

Figure 1 shows a schematic representation of the layout of the sensor suite and the beam splitters that are deployed to direct the appropriate band of the incoming radiation to each of the 3 individual sensors. The incoming radiation is first split into a longwave (thermal) and a visual+NIR part by a heat reflecting (hot) mirror (a custom made Melles Griot dichroic beam splitter consisting of Schott N-BK7 Borosilicate Crown glass with an Indium Tin Oxide coating, with a reflection  $R > 85\%$ ). The longwave part of the spectrum is reflected into the lens of the thermal camera, while the visual+NIR light is transmitted to a combination of two digital image intensifiers that are mounted under an angle of 90 degrees. Next, a near-infrared reflecting mirror (45 deg angle of incidence, Borofloat glass, type Edmund Optics B43-958, 101x127x3.3 mm, see: [www.edmundoptics.com](http://www.edmundoptics.com)) is used to separate the incoming light, by transmitting the visual (400–700nm) and reflecting the NIR part (700–900nm), such that one image intensifier registers the visual part and the other one only detects the NIR part of the incoming radiation. The sensor geometry is such that the optical axes of all cameras are aligned.

The two image intensifiers are high resolution (1280x960) Photonis PP3000U Intensified Camera Units (ICUs, Fig. 2a: [www.photonis.com](http://www.photonis.com)). The ICU is a low light level, intensified CMOS camera. It has a 2/3" CMOS sensor with a spectral response range of 400-900 nm, and delivers both a PAL or NTSC composite video signal output (ITU-R BT.656-4, 640x480 pixels), and an SDI – LVDS 270 Mbits/s signal. Both ICU's are equipped with Pentax C2514M CCTV lenses, with a minimal focal length of 25mm and a lens aperture of F/1.4, resulting in a FOV of 30.7° x 24.8°.

The thermal camera is a XenICs Gobi 384 uncooled a-Si infrared microbolometer (Fig. 2b: [www.xenics.com](http://www.xenics.com)). It has a 384 x 288 pixel focal plane array, and a spectral sensitivity range of 8–14μm, which is the range of most interest for outdoor applications. It is equipped with an Ophir supIR18mm F/1 lens ([www.ophiropt.com](http://www.ophiropt.com)) providing a 29.9° x 22.6° wide angle view. The Gobi 384 has a 16-bit Ethernet and CameraLink interface.

The sensors and the mirrors are mounted on a common metal base. The whole configuration is placed in an enclosed housing.

The sensor suite delivers both analog video and digital signal output.

### 2.2. GPS Receivers

An internal U-blox EVK-5P Positioning Engine ([www.u-blox.com](http://www.u-blox.com)) provides a position and orientation (i.e. sensor location and viewing direction) signal through the high-speed 7-port USB 2.0 hub. The accuracy in position is less than 3m. The accuracy in orientation is less than 5 degrees.

In local area operations, and when high accuracy is needed, an external Trimble SPS751 GPS receiver set ([www.trimble.com](http://www.trimble.com)) is connected to the system, to achieve high position accuracy (< 1 cm) through real time kinematic (RTK) GPS signal correction.

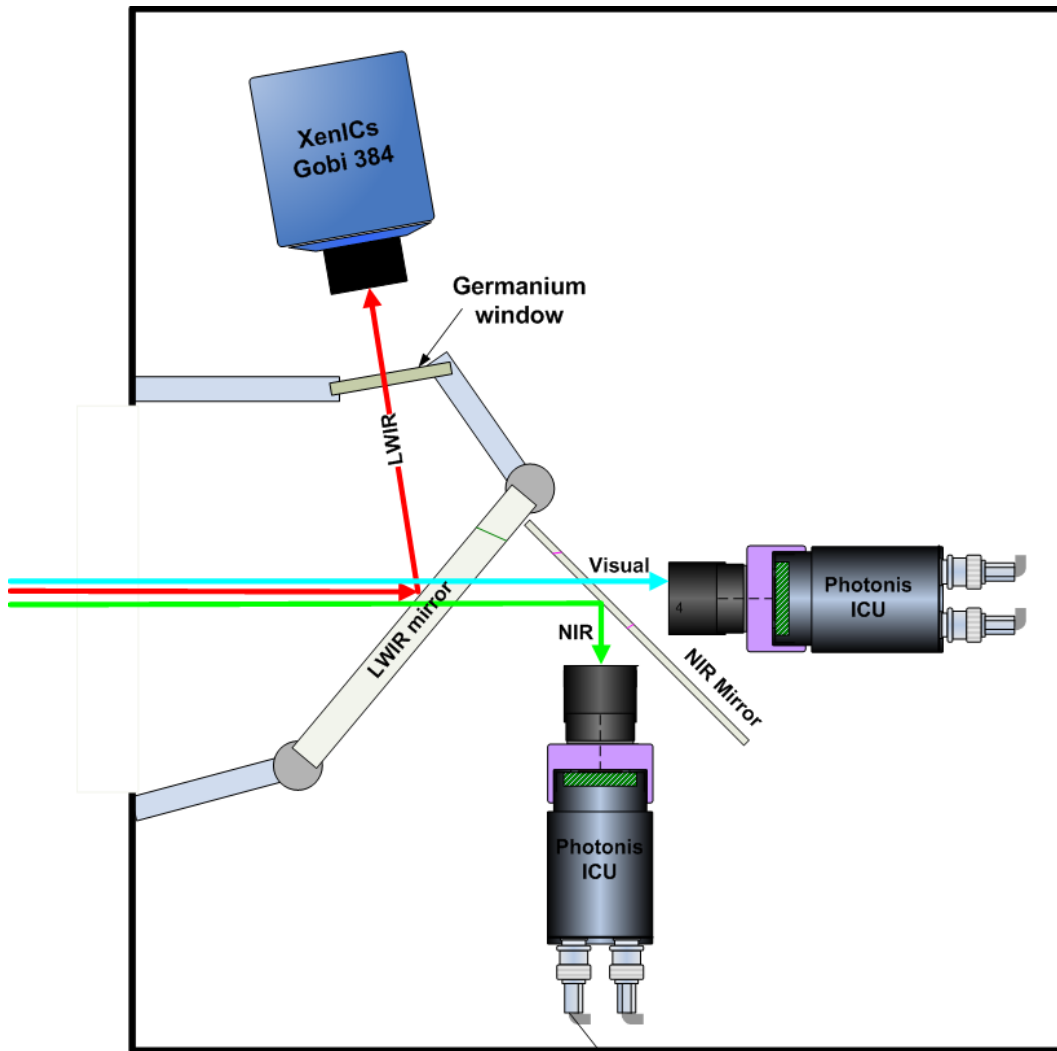


Fig. 1. Layout of the sensors and filters of the INVIS sensor suite.



(a)



(b)

Fig. 2. (a) Photonis PP3000U Intensified Camera Unit, and (b) XenICs Gobi 384 microbolometer.



(a)



(b)



(c)

Fig. 3. The INVIS sensor suite. (a) Top view, (b) inside , and (c) the sensor suite mounted on a mobile all-terrain platform.

### 2.3. Electronic Compasses

An internal Silicon Labs F350-COMPASS-RD multi-axis electronic compass ([www.silabs.com](http://www.silabs.com)) provides the azimuth and tilt angle of the optical axis of the sensor suite, with an accuracy of a few degrees.

When the viewing direction needs to be known with higher accuracy, an external Xsens 3D inertial measurement unit (IMU) motion sensor with accelerometer, magnetometer and gyroscope ([www.xsens.com](http://www.xsens.com)) is connected to the system to measure Yaw, Roll en Pitch with an accuracy less than  $0.1^\circ$ .

### 2.4. Computer

A Dell Precision M2400 Intel Core Duo P8600 2.4GHz laptop with a solid state harddisk is used to store, colorize, and visualize the sensor signals and to generate and display the synthetic scene view. The current implementation achieves real-time ( $\sim 25$  Hz) visualization, enhancement and data registration.

## 2.5. Displays

Two 6.4" TFT video displays, embedded in the system casing, enable simultaneous monitoring of two of the three video signals (either Visual/NIR, Visual/LWIR, or NIR/LWIR). The laptop display (14 inch, 1440x900 pixels) is used to view the final fused, colored and enhanced images, combined with the synthetic 3D scene views.

## 2.6. Data Transfer and Storage

The Photonis ICU's are connected to a high-speed 7-port USB 2.0 hub. This enables the user to interface with the ICU's and to adjust their settings, or to download and install preferred settings.

A Pleora iPORT PT1000-ANL-2/6 frame grabber ([www.pleora.com](http://www.pleora.com)) is used to digitize the analog video output signals of (1) both ICU's and (2) the Gobi 384. Digitization is performed at video rate, with at a resolution of 640x480 pixels and 10 bits per pixel. The Pleora transmits these signals to a Netgear Gigabit Ethernet switch. The 16-bit TCP/IP Ethernet interface of the XenICs Gobi 384 is also directly connected to the Netgear Gigabit Ethernet switch.

Three Pinnacle Video Transfer Units ([www.pinnaclesys.com/PVT](http://www.pinnaclesys.com/PVT)) are provided to store (a) the analog video signals of all three cameras, and (b) the audio signals of two (optional) external microphones, either on 3 internal 320 Gb harddisks, or on USB memory sticks. The microphones can for instance be positioned on the front and back of the camera suite. The microphone on front can then be used to register relevant audio information from the registered scene, and the second microphone can for instance be used to record spoken annotations.

## 3. SYSTEM SOFTWARE

The INVIS system software comprises (a) a color remapping procedure, (b) several image enhancement procedures, and (c) software to generate a synthetic image of the environment. Color remapping serves to represent the fused sensor signals in natural daytime colors, which makes nighttime imagery easier to interpret (more intuitive)<sup>3</sup>. It can also be used to selectively enhance target conspicuity by increasing target color contrast. Image enhancement serves to increase perceptual image quality, by reducing the amount of noise, increasing the dynamic range, and increasing image resolution. Augmenting nighttime imagery with synthetic scene views serves to enhance situational awareness.

### 3.1. Color Mapping

The signals from the three nighttime cameras are fused and represented in natural daytime colors using the Color-the-Night false color remapping procedure<sup>10,11</sup>. Since the color mapping operation is efficiently implemented as a lookup table transform<sup>2,4,5</sup>, only a minimal amount of computation is required to achieve real-time performance. Color fusion can therefore either be performed on a standard PC. For reasons of portability we chose to use a standard laptop, connected to the sensor suite via an Ethernet connection.

The principle of the lookup-table based color mapping technique is explained in detail elsewhere<sup>2,4,5</sup>. For the sake of completeness we will now briefly describe this procedure. First, the LWIR, Visual and NIR bands of the INVIS sensor suite are mapped to respectively the R, G, and B channels of a false color RGB image. Second, the false color image thus obtained is transformed into an indexed image, with a corresponding 3D lookup table consisting of a set of RGB triples. Finally, the lookup table corresponding to the input false color multiband nightvision image is replaced (swapped) with a new color lookup table that maps the three image bands onto natural colors. The new color lookup-table can be obtained either by applying a statistical transform to the entries of the original lookup-table, or by a procedure that replaces entries of the original lookup-table by their corresponding natural color values. The *statistical transform method* transfers the first order statistics (mean and standard deviation) of the color distribution of a representative natural color daytime reference image to the false color multiband nighttime image<sup>4,11</sup>. This mapping is usually performed in a perceptually decorrelated color space (e.g.  $lab^*$ ). The *sample-based method* deploys a set of corresponding samples from the combination of a multi-band sensor image of a given scene and a registered naturally colored (RGB) daytime reference image of the same scene, to derive a color lookup table transform pair that transfers the color characteristics of the

natural color reference image to the false-color nighttime image<sup>2,4</sup>. For an 8-bit three-band system the 3D color lookup table contains 256x256x256 entries. When the color lookup table contains fewer entries, the color mapping is achieved by determining the closest match of the table entries to the observed multi-band sensor values. Once the color transformation has been derived and the color lookup table pair that defines the mapping has been created, the table pair remains fixed and can be deployed in a real-time application. The lookup table transform requires minimal computing power. An additional advantage of the color lookup transform method is that object colors only depend on the multi-band sensor values, and are independent of the image content. As a result, objects keep the same color over time when registered with a moving camera.

## 3.2. Image Enhancement

### 3.2.1. Dynamic Super Resolution (DSR)

Camera images suffer from temporal noise, which may hinder target detection and comfortable observation. Noise reduction serves to reduce temporal noise. By combining information from several frames, details can be enhanced. This is done by using the Resolution Enhancement algorithm. The Dynamic Super Resolution (DSR) algorithm<sup>9</sup> can both be used for Resolution Enhancement in combination with Noise Reduction, or for Noise Reduction only. For optimal performance it relies on accurate motion estimation, which is provided by a Scene Based Motion Estimation (SBME) algorithm. SBME estimates image motion from the image information itself.

For intensified imagery only noise reduction is applied, since these images are mostly degraded by noise. The thermal camera image contains more aliasing, in which case resolution enhancement will produce better results.

### 3.2.2. Scene-Based Non-Uniformity Correction (SBNUC)

Scene-Based Non-Uniformity Correction (SBNUC) serves to correct for the different response characteristics of the detectors in the thermal camera sensor<sup>9</sup>. SBNUC is based on the assumption that the scene content moves slowly relative to the frame rate, whereas the non-uniformities are semi-static. Accurate motion estimation is used to identify the actual apparent scene movement, from which a motion-corrected difference image can be calculated, containing the current residual non-uniformity with some additional temporal noise. Based upon an advanced image formation model, SBNUC provides optimal reduction of the non-uniformities over time.

### 3.2.3. Local Adaptive Contrast Enhancement (LACE)

Local Adaptive Contrast Enhancement (LACE) is a multiresolution technique to compress the overall dynamic range of the images, such that small amplitude details are retained and an overall natural look is maintained<sup>8</sup>.

For an input image  $I$ , the enhanced output image  $O$  is given by

$$O = I + \sum_{i=1}^k c(\sigma_i) \cdot (I - \mu_i)$$

where  $\mu_i$  and  $\sigma_i$  represent respectively the local mean and standard deviation of the input image at resolution level  $i$ , and the multiplication factor is given by

$$c(\sigma) = \alpha \frac{M}{\sigma} - 1$$

where  $M$  is the global mean of  $I$ , and  $\alpha$  is a free parameter.  $c$  is clipped

LACE has 3 free parameters: the number of resolution levels ( $k$ ), the weight of a processed level relative to the original image, and the maximum allowed multiplication<sup>8</sup>.

Since LACE alters the dynamic range considerably, it is applied after the image fusion and color remapping procedures.

### 3.3. Synthetic Scene Generation

The INVIS system is capable of displaying synthetic 3-D views of the environment. It combines a priori information about the area of operation with real-time input from the position information system, to produce a corresponding view of the environment. The system augments the Triclob's<sup>12</sup> sensory views to increase and maintain the operator's situation awareness. The system is particularly beneficial in adverse real-world viewing conditions (e.g. due to bad atmospheric conditions) or when the operator's view is obstructed by an obstacle.

The environment information used by the synthetic vision system consists of a number of components:

- A 3-D visual representation of the area of operation, containing terrain elevation, buildings and their interiors, vegetation, and other static or more or less permanent objects in the environment.
- A classification of the objects in the environment into a discrete number of types (i.e. buildings, building walls and doors, vegetation, transportation surfaces, et cetera).
- Material classifications for all surfaces in the environment.

The synthetic vision system is built using TNO's Enhanced Virtual Environment (EVE), a data-driven and data-centric framework for 3-D simulation and visualization. The system consists of a 3-D view and a number of other views that are capable of displaying sensor images. On top of that, a graphical user interface has been constructed. This GUI allows the user to control the visual appearance of the 3-D virtual environment and to control the timing of a scenario, in the case of a replay of data that was recorded earlier.

A geo-specific 3-D visual terrain database of the Marnehuizen MOUT training village has been used for experiments. This database has been modeled using construction blueprints and close-up photographs. Because of possible differences between the original blueprints and the final construction, the accuracy of the terrain database is estimated at 1 meter.

The synthetic vision system offers a number of features that allow the operator to manipulate his view on the world and acquire insight in his situation.

#### 3.3.1. Play Modes

The system is capable of visualizing the virtual environment in two discrete play modes. The first mode is the live-play mode, in which the system uses up-to-date information to display the virtual environment in real-time. The second mode is the replay mode, in which recorded data (both position and information data as well as sensor images) is retrieved from a database. In replay mode, the user is able to control the timing of the scenario by means of a time slider bar and a number of buttons (e.g. reverse, forward, fast-forward).

#### 3.3.2. View Modes

The system provides a discrete number of view modes. These are views from first person and third person perspective, and a free-look perspective. The first-person view uses recorded position and orientation information to provide an accurate view of the synthetic environment from the sensor's perspective. From the first-person view, the operator is able to directly match the information visible on the sensory views with the information from the synthetic view. This information provides cues for object recognition and classification.



Fig. 4. Corresponding LWIR (left) and synthetic (right) images in a first-person perspective.



The third-person view allows the operator to view the environment from a perspective that is relative and attached to the sensor's actual position. In this view, the sensor is represented visually in the environment to provide cues about the sensor's position and orientation in the area of operation (Fig.5b). Using this view, it is relatively easy for the operator to acquire insight into otherwise difficult-to-view parts of the area of operation, while at the same time keeping track of his own position and local surroundings.



Fig. 5. Corresponding LWIR (a) and synthetic (b) images from third-person perspective, with the sensor's view cone visible in the synthetic view

The free-look view is similar to the third-person view, in the sense that it is also external. The free-look view, however, is not attached to the sensor's actual position and allows the operator to freely explore the environment. Compared to the third-person view, this view involves more freedom when it comes to exploration of the area of operation, while at the same time possibly losing track of the operator's own position.

### 3.3.3. Object Class Switches

From any view in the synthetic vision system, the operator is able to "switch off" classes of objects. This feature is particularly effective when both sensory and synthetic views are densely cluttered. Also, this feature allows the user to see through obstacles. This can be employed to quickly look behind a certain object, by removing it entirely from the view, or to view an object's interior (Fig.6-7c), by removing it partially from the view. This feature can be employed for the planning of possible paths to traverse or the identification of potential hiding places for opposing forces.

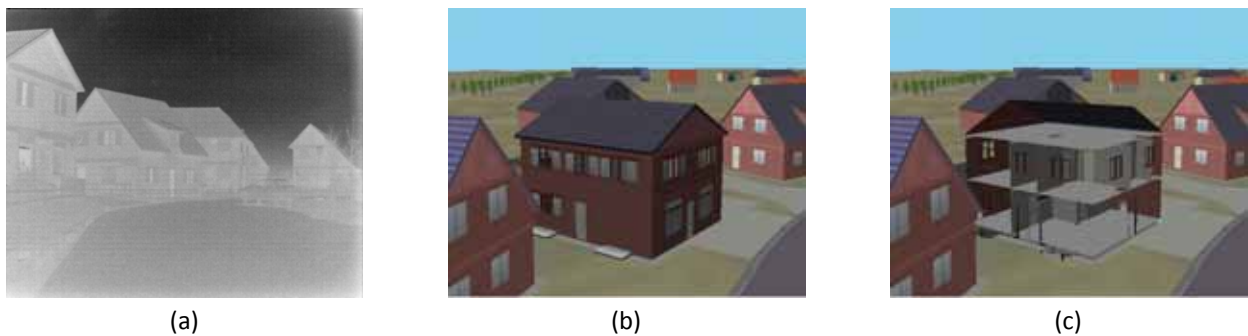


Fig. 6. (a) LWIR view of an environment with a building, and synthetic views of the same building with walls switched on (b) and off (c), seen from a different perspective.





Fig. 7. (a) LWIR view of the environment, synthetic views with vegetation switched on (b) and off (c).

#### 3.3.4. Material View

The material view switches texture mapping of the synthetic environment's surfaces. In this view, the original photorealistic textures are replaced by color textures that indicate material characteristics. Aspects such as material type and trafficability of ground surface elements can therefore be distinguished easily by the operator and used for object type classification and navigation purposes.

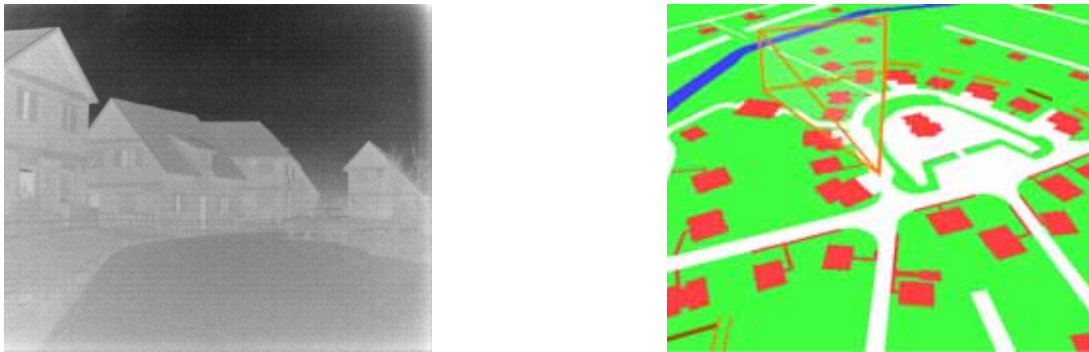


Fig. 8. LWIR view (left) with corresponding material view from third-person perspective (right).

#### 3.3.5. Integration with the Triclobes system

The synthetic vision system is integrated into the Triclobes<sup>12</sup> system in several ways, both using information gathered by the system and providing information back. In this paragraph, the interfaces between the synthetic vision system and other parts of the Triclobes system are described.

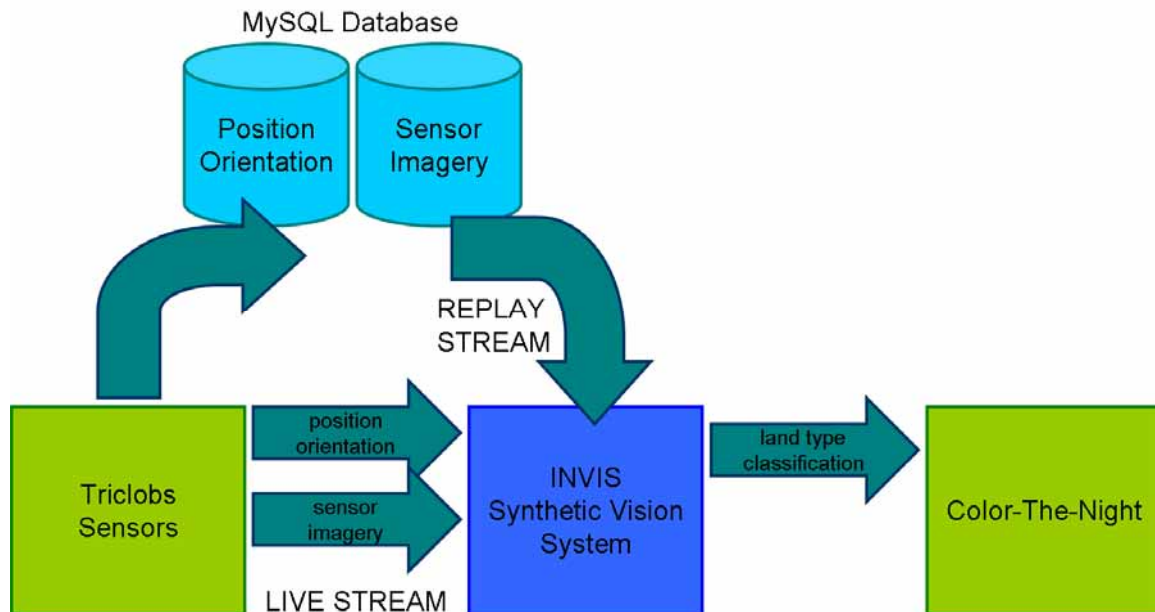


Fig. 9. System overview. The arrows represent the interfaces to and from the synthetic vision system.

### 3.3.6. Position Information System Interface

As described earlier, the synthetic vision system uses information from Triclobes' position and orientation sensors to pinpoint the sensor's location within the synthetic environment. The system is currently capable to receive information both through a dedicated interface and through a standardized NMEA interface. In replay mode, this data is retrieved from a MySQL database.

### 3.3.7. Night Vision Sensor Interface

Besides displaying synthetic views, the synthetic vision system is also capable of directly showing sensor images. These images can be displayed either next to the synthetic image in a 4x4 tiled window, or as image overlays on top of the synthetic image. The latter view mode allows for precise visual correlation between sensor images and the synthetic, first-person image. In replay mode, this data is retrieved from a MySQL database.

### 3.3.8. Color-The-Night Interface

The Color-The-Night algorithm uses an environment-specific mapping (e.g. urban, countryside, or forest) to create synthetic daylight images from sensor images. Knowledge about the specific type of environment that the sensor is observing can therefore be used as a cue for choosing the correct mapping. By reasoning about the content of the visible environment within the synthetic vision system's field of view, it is possible to deduce which type of land is observed.

Based on a configurable mapping table, the system is now capable of mapping the virtual environment's textures to land type classifications. By counting all the pixels in screen that belong to a certain land type, a histogram of the distribution of land types is eventually calculated. These histograms are then matched to pre-calculated land type histograms that belong to specific Color-The-Night mapping schemes. Using a least-square matching algorithm, the best match is selected as the Color-The-Night mapping scheme used for the current frame.

The Color-The-Night interface is currently in an experimental phase. Further experimentation and optimization are required to improve the mapping deduction algorithm and to make this interface effective for real-time use.

#### 4. FIELD TRIALS

We tested the INVIS system during two nocturnal data collection trials in the field. The first trial was at a simulated typical Bosnian area at the camouflage school in Reek, The Netherlands (Fig. 10). The second trial was in Marnehuizen, a Dutch MOUT village (see Figs. 11-13; Google Earth: Latitude 53.386293 and Longitude 6.262944). At both occasions ambient light levels were below 0.01 lx.

Figure 10d shows a false color representation of a typical rural Bosnian scene, obtained by mapping the visual (Fig.10a), NIR (Fig.10b), and LWIR (Fig.10c) bands to respectively the R, G, and B color channels. The natural looking color image in Figure 10f is obtained by applying Color-the-Night remapping to Figure 10d, using a color mapping derived from a corresponding daylight color photograph (Fig.10e). The perceptual quality of this image is further enhanced by dynamic noise reduction (Fig.10g) and local contrast enhancement (Fig.10h).

Figures 10-13 (d) show false color representations of typical scenes from the MOUT village Marnehuizen, obtained by mapping the visual (a), NIR (b), and LWIR (c) bands to respectively the R, G, and B color channels. The natural looking color images Figs. 10-13 (f) are obtained by applying Color-the-Night remapping to Figs. 10-13(d). Note that the reference daylight color image from which the color mapping was derived is the same in all cases (Figs.10-13e). This illustrates the robustness of the color mapping procedure: object colors are independent of the scene content, and once the mapping has been derived it can be used in similar environments. The perceptual quality of these images is further enhanced by dynamic noise reduction (Figs.10-13g) and local contrast enhancement (Figs.10-13h).

Figures 10-13 (j) show that color nighttime imagery with a realistic daytime color appearance can also be obtained by using color mapping schemes derived from synthetic scene views (i). This is a practical feature when no daylight color imagery is available for certain areas. Figures 10-13 (k) and (l) show these synthetically colored nighttime images after respectively dynamic noise reduction (k) and local contrast enhancement (l).

Note that the left part of the scene in Figure 13 is occluded by smoke from a grenade in the visual (a) and NIR (b) bands, but not in the LWIR band (c). This scene demonstrates that sensor fusion makes the system robust against sensor drop out.

#### 5. CONCLUSIONS

In this paper we presented the prototype INVIS portable surveillance and observation system. The INVIS system provides real-time co-aligned visual, near-infrared and thermal images, augmented with synthetic scene views. All images can either be stored on on-board harddisks, and can be enhanced in real-time by a laptop computer. A real-time color remapping procedure, implemented as a lookup table transform, is used to fuse and represent the three camera signals in natural daytime colors. Synthetic scene views are generated in real-time using a geometric 3D scene model in combination with the position and orientation information provided by the GPS and inertial sensors of the INVIS system. The results of some preliminary field trials clearly demonstrate the benefits of this system for surveillance, navigation and target detection tasks. The resulting false color nightvision images closely resemble daytime color images, while thermal targets are clearly distinguishable. The synthetic scene view enhances the user's situational awareness by providing information that is not available from each of the individual imaging sensors.

In a later stage we plan to implement the option to visualize the synthetic scene views either as overlays on top of, or fused with, the dynamic nighttime imagery<sup>13</sup>, to produce an augmented reality representation of the scene<sup>1</sup>. This will provide valuable positional information when objects are occluded (e.g. by dust from explosions, vegetation, or fog), or structural information on (partly) destroyed objects (e.g. collapsed buildings). Additional information like the position of electricity cables and gas or water pipes can also be superimposed on the scene being viewed. Our final goal is to produce an augmented virtual representation of the environment<sup>6</sup>.

#### Acknowledgement

Effort sponsored by the Air Force Office of Scientific Research, Air Force Material Command, USAF, under grant number FA8655-09-3095. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purpose notwithstanding any copyright notation thereon.

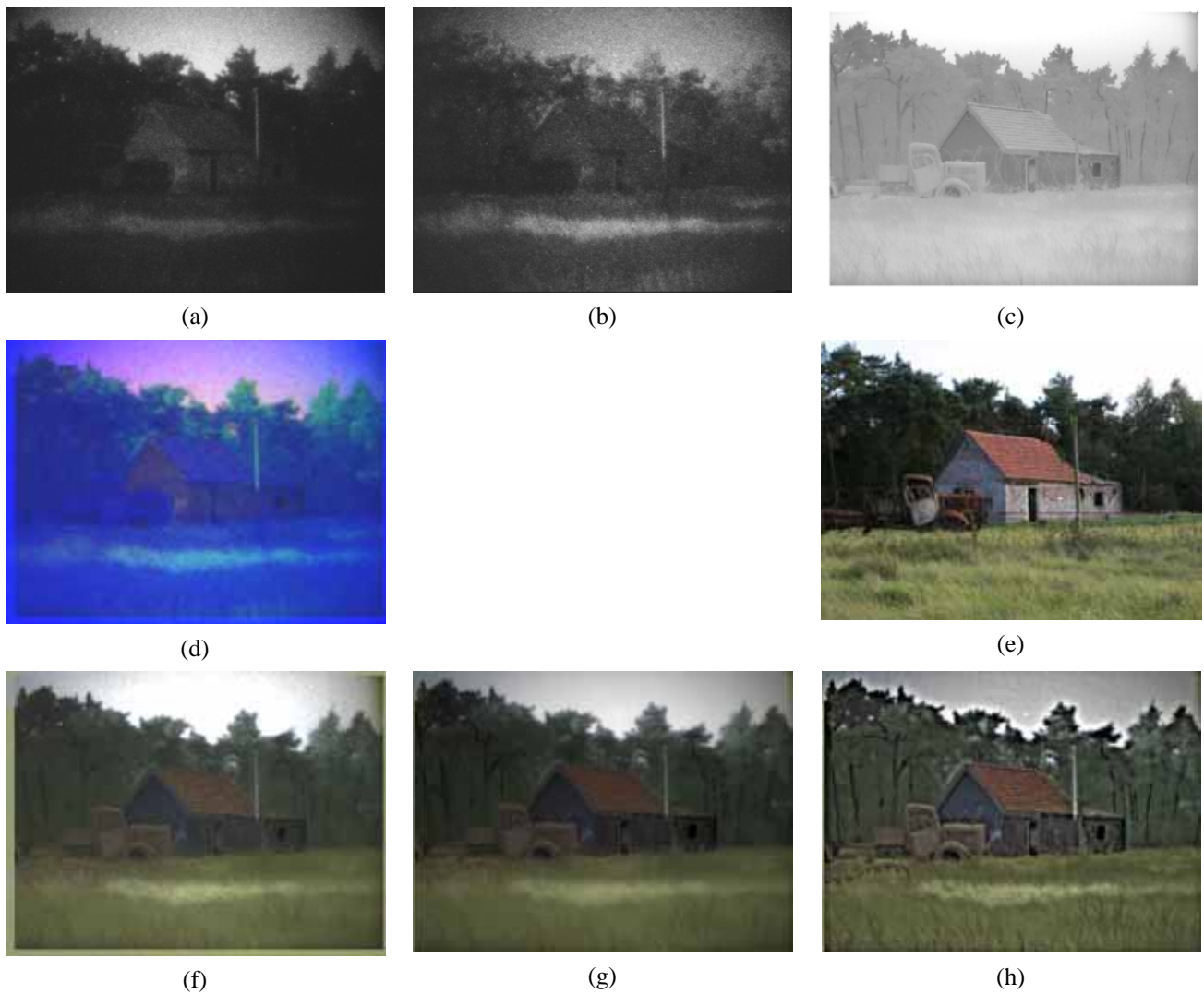


Fig. 10. (a) Visual, (b) NIR and (c) LWIR input signals. (d) False color image obtained by mapping the three bands to respectively the R, G, and B channels. (e) Full color daytime reference image. (f) Image obtained after applying Color-the-Night remapping to the image shown in (d), using a mapping derived from (e). (g) The image shown in (f) after dynamic noise reduction. (h) Image from (g) after local contrast enhancement.



Fig. 11. Top-down, left to right: (a-c) Visual, NIR and LWIR input signals. (d) False color image obtained by mapping the three bands to respectively the R, G, and B channels. (e) Full color daytime reference image. (f) Image obtained after applying Color-the-Night remapping to the image shown in (d), using a mapping derived from (e). (g) The image shown in (f) after dynamic noise reduction. (h) Image from (g) after local contrast enhancement. (i) Synthetic view of the scene depicted in (d). (j-l) As (f-h) after applying Color-the-Night remapping to the image shown in (d), using a mapping derived from the synthetic scene (i).



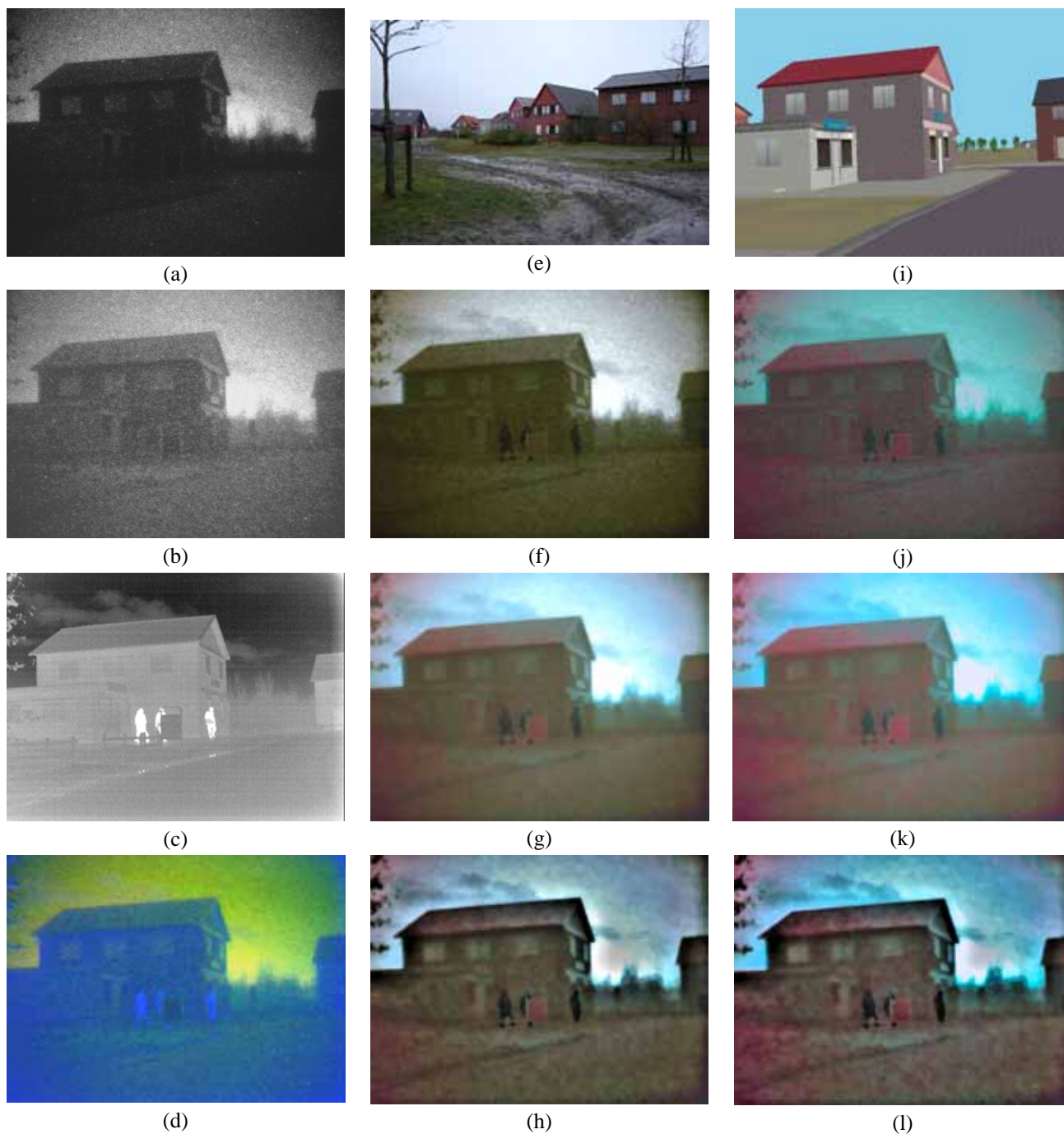


Fig. 12. Top-down, left to right: (a-c) Visual, NIR and LWIR input signals. (d) False color image obtained by mapping the three bands to respectively the R, G, and B channels. (e) Full color daytime reference image. (f) Image obtained after applying Color-the-Night remapping to the image shown in (d), using a mapping derived from (e). (g) The image shown in (f) after dynamic noise reduction. (h) Image from (g) after local contrast enhancement. (i) Synthetic view of the scene depicted in (d). (j-l) As (f-h) after applying Color-the-Night remapping to the image shown in (d), using a mapping derived from the synthetic scene (i).



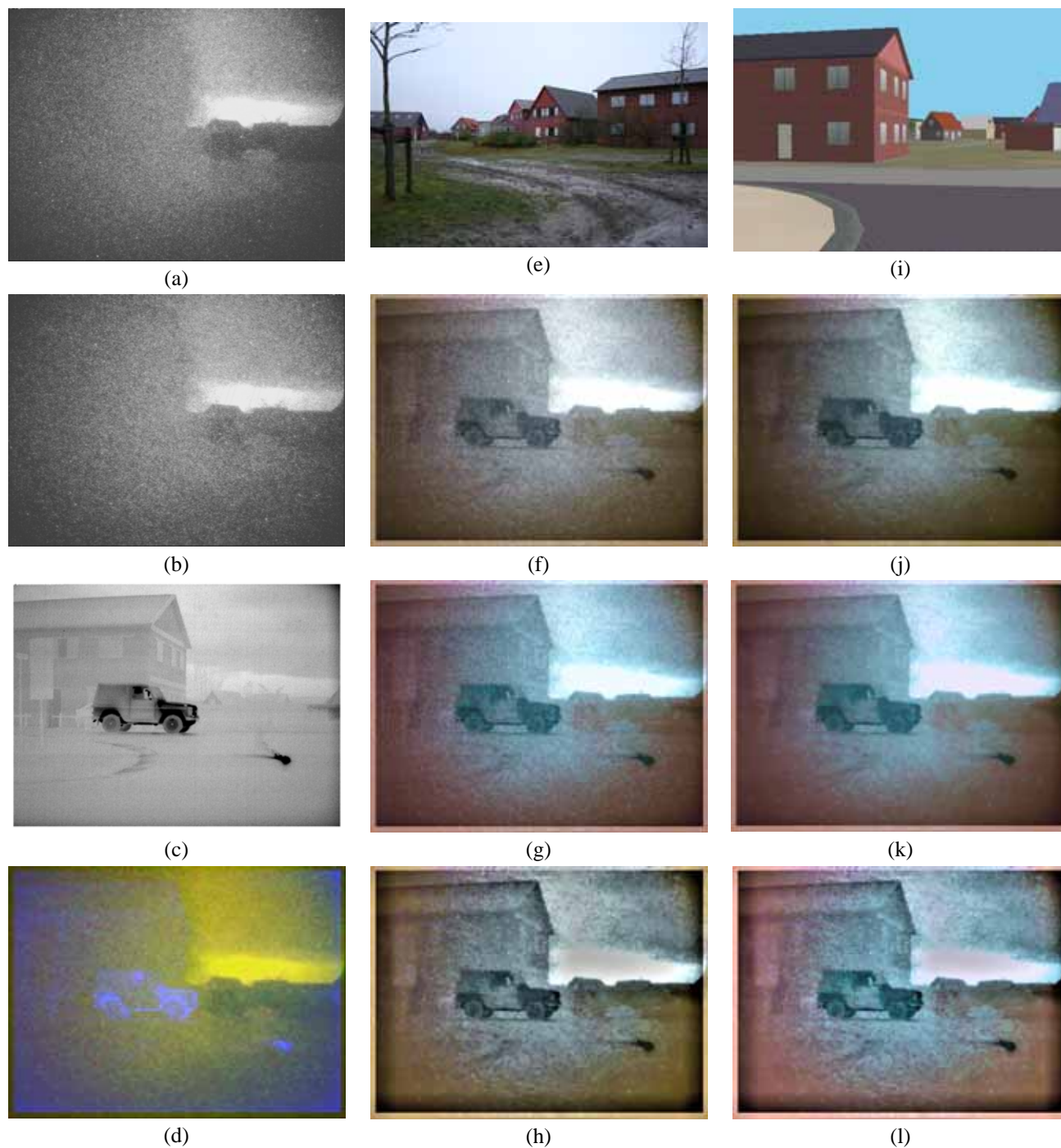


Fig. 13. Top-down, left to right: (a-c) Visual, NIR and LWIR input signals. (d) False color image obtained by mapping the three bands to respectively the R, G, and B channels. (e) Full color daytime reference image. (f) Image obtained after applying Color-the-Night remapping to the image shown in (d), using a mapping derived from (e). (g) The image shown in (f) after dynamic noise reduction. (h) Image from (g) after local contrast enhancement. (i) Synthetic view of the scene depicted in (d). (j-l) As (f-h) after applying Color-the-Night remapping to the image shown in (d), using a mapping derived from the synthetic scene (i). Note that the left part of the scene is occluded by smoke from a grenade in the visual (a) and NIR (b) bands, but not in the LWIR band (c).

## REFERENCES

1. Azuma, R.T., A survey of augmented reality, *Presence: Tele-operators and Virtual environments*, 6(4) ,pp. 355-385, 1997.
2. Hogervorst, M.A. and Toet, A., Method for applying daytime colors to nighttime imagery in realtime, In: B.V. Dasarathy (Ed.), *Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications 2008*, pp. 697403-1-697403-9, The International Society for Optical Engineering, Bellingham, WA, USA, 2008.
3. Hogervorst, M.A. and Toet, A., Evaluation of a color fused dual-band NVG-sensor, In: B.V. Dasarathy (Ed.), *Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications 2009*, pp. 734502-1-734502-7, SPIE - The International Society for Optical Engineering, Bellingham, WA, 2009.
4. Hogervorst, M.A. and Toet, A., Fast natural color mapping for night-time imagery, *Information Fusion*, 11(2) ,pp. 69-77, 2010.
5. Hogervorst, M.A., Toet, A., & Kooi, F.L. (2006). TNO Defense Security and Safety. Method and system for converting at least one first-spectrum image into a second-spectrum image. Patent Number PCT/NL2007050392. Application Number 0780855.5-2202,
6. Neumann, U., You, S., Hu, J., Jiang, B. and Sebe, I.O., Visualizing reality in an augmented virtual environment, *Presence: Tele-operators and Virtual environments*, 13(2) ,pp. 222-233, 2004.
7. Ruderman, D.L., Cronin, T.W. and Chiao, C.-C., Statistics of cone responses to natural images: implications for visual coding, *Journal of the Optical Society of America A*, 15(8) ,pp. 2036-2045, 1998.
8. Schutte, K., Multi-scale adaptive gain control of IR images, In: B.F. Andresen & M. Strojnik (Ed.), *Infrared Technology and Applications XXIII*, pp. 906-914, The International Society for Optical Engineering, Bellingham, WA, 1997.
9. Schutte, K., de Lange, D.J. and van den Broek, S.P., Signal conditioning algorithms for enhanced tactical sensor imagery, In: G.C. Holst (Ed.), *Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XIV*, pp. 92-100, The International Society for Optical Engineering, Bellingham, WA, 2003.
10. Toet, A., Color the night: applying daytime colors to nighttime imagery, In: J.G. Verly (Ed.), *Enhanced and Synthetic Vision 2003*, pp. 168-178, The International Society for Optical Engineering, Bellingham, WA, USA, 2003.
11. Toet, A., Natural colour mapping for multiband nightvision imagery, *Information Fusion*, 4(3) ,pp. 155-166, 2003.
12. Toet, A. and Hogervorst, M.A., The TRICLOBS portable triband lowlight color observation system, In: B.V. Dasarathy (Ed.), *Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications 2009*, pp. 734503-1-734503-11, SPIE - The International Society for Optical Engineering, Bellingham, WA, 2009.
13. You, S. and Neumann, U., Fusion of vision and gyro tracking for robust augmented reality registration, In: *Proceedings of the IEEE Virtual Reality Conference 2001 (VR 2001)*, pp. 7-79, IEEE Press, Washington, USA, 2001.