Literature Review 1 – Computer Graphics

Viswanath Pulle, 01690370

Primary Paper: *A Virtual Try-On System for Prescription Eyeglasses*

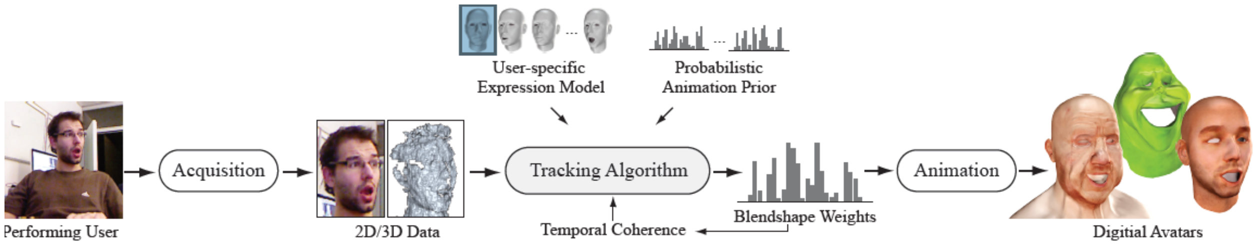
Secondary Paper: *Real time Performance-Based Facial Animation*

This primary paper addresses a solution to a problem for people trying on new prescription eyeglasses with virtual try on feature in the online stores. They are showcased with demo lenses, hence ignoring the refractive nature of the lenses and hence do not show the exact appearance of the final glasses. It makes a huge difference when we are dealing with high prescriptions and could significantly change the look of a person. This paper addresses the refractive properties of the lenses delivering more accurate simulations before making a purchase. The paper describes using image techniques to generate a 3D representation of the corrective lenses mounted into the eyeglasses frame and modifies the video sequence to virtually insert the eyeglasses using image-based rendering. Eventually, giving users a better idea of how they would look when wearing the new pair of eyeglasses.

The authors used another paper (My secondary paper in discussion) for reference to make use of the performance-based character animation that enables any user to control the facial expressions of a digital avatar in real time. This technology to map facial movements to a digital image have been used by the primary paper authors to simulate the existence of eyeglasses on the user's face virtually. The paper introduces a novel face tracking algorithm that combines geometry and texture registration with pre -recorded animation priors in a single optimization. It is formulated as a maximum a posteriori estimation in a reduced parameter space. This makes their system easy to deploy and facilitates a range of new applications, including the Primary authors' application of virtual try on Eyeglasses.

The main contribution of this paper is the face tracking algorithm that combines 3D geometry and 2D texture registration in a systematic way with dynamic blend shape priors generated from existing face animation sequences. Formulated as a probabilistic optimization problem, our method successfully tracks complex facial expressions even for very noisy inputs. This is achieved by mapping the acquired depth maps and images of the performing user into the space of realistic facial expressions defined by the animation prior. Real time processing is facilitated by a reduced facial expression model that can be easily adapted to the specific expression space and facial geometry of different users. We integrate these components into a complete framework for real time, non-intrusive, marker less facial performance capture and animation.

Performance-driven facial animation requires solving two main technical challenges: We need to accurately track the rigid and non-rigid motion of the user's face, and map the extracted tracking parameters to suitable animation controls that drive the virtual character. Our approach combines these two problems into a single optimization that solves for the most likely

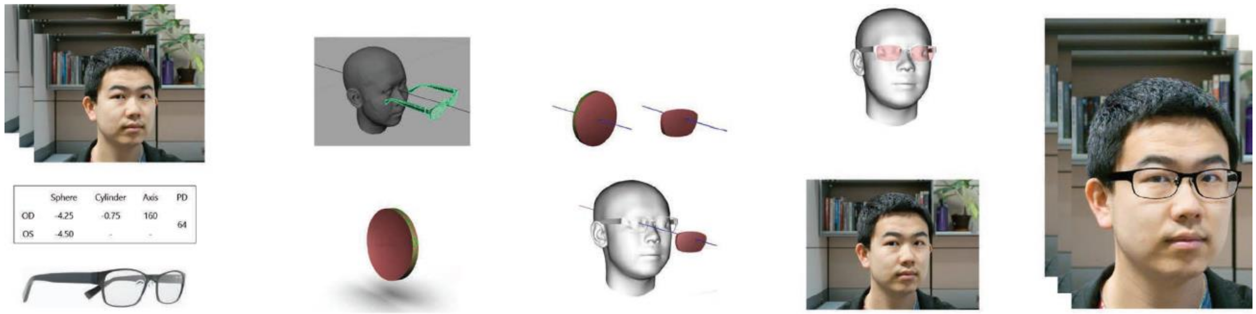


parameters of a user specific expression model given the observed 2D and 3D data. We derive a suitable probabilistic prior for this optimization from prerecorded animation sequences that define the space of realistic facial expressions. The figure above gives an overview of the pipeline of this paper.

Blend shape Representation: To integrate tracking and animation into one optimization, we represent facial expressions as a weighted sum of blend shape meshes. This design choice offers a number of advantages: A blend shape model provides a compact representation of the facial expression space, thus significantly reducing the dimensionality of the optimization problem. In addition, we can use existing blend shape animations, that are ubiquitous in movie and game production, to define the dynamic expression priors.
The underlying hypothesis here is that the blend shape weights of a human facial animation sequence provide a sufficient level of abstraction to enable expression transfer between different characters. Finally, the output generated by our algorithm, a temporal sequence of blendshape weights, can be directly imported into commercial animation tools, thus facilitating integration into existing production workflows.
All input data is acquired using the Kinect system, i.e. no other hardware such as laser scanners is required for user-specific model building. The Kinect supports simultaneous capture of a 2D color image and a 3D depth map at 30 frames per second, based on invisible infrared projection. Essential benefits of this low-cost acquisition device include ease of deployment and sustained operability in a natural environment. The user is neither required to wear any physical markers or specialized makeup, nor is the performance adversely affected by intrusive light projections or clumsy hardware contraptions. However, these key advantages come at the price of a substantial degradation in data quality compared to state-of-the-art performance capture systems based on markers and/or active lighting. Ensuring robust processing given the low resolution and high noise levels of the input data is the primary challenge that we need to be address.

The virtual try-on system inserts prescription eyeglasses onto the user's face and simulates important changes to the appearance due to refraction, reflection, or shadows cast on the face. The system uses the following three elements as input:
Image sequence: An image sequence of the user without eyeglasses is captured with a color camera
User's eyeglasses prescription: An eyeglasses prescription, usually provided by an optometrist, and specifies the value of all parameters necessary to correct blurred vision due to refractive errors, like myopia, hyperopia, presbyopia, and astigmatism.
Eyeglasses frame: The user chooses the desired eyeglasses frame.

The figure below gives an overview of the paper's pipeline, which consists of two main stages, virtual eyeglasses generation and video synthesis:



Virtual Eyeglasses Generation:
In this stage, a 3D representation of the prescription eyeglasses (the frame and corrective lenses), with an appropriate position relative to the user’s face geometry. Inspired by the traditional eyeglasses manufacturing pipeline, this virtual eyeglasses generation stage has three steps:

1. *Positioning the eyeglasses on the user’s face*. After an initial manual positioning step for the first frame, we use face tracking to automatically align the eyeglasses with the user’s face in the following frames.
2. *Creating a parametric lens model.* Based on the user’s prescription and desired lens properties, this model describes the geometry of the uncut lens before mounting.
3. *Cutting and mounting the lens.* We trim the lens geometry according to the shape of the eyeglasses frame and insert the virtual lenses into the eyeglasses frame.

Video synthesis stage:
In this stage we render the virtual eyeglasses and insert them into the input image sequences, taking into account the eyeglasses frame, lenses, and surrounding lighting. In doing so, we account for the effects of refraction, reflection, and shadows due to the inserted eyeglasses. we first place the eyeglasses frame with respect to the user’s face geometry.
We obtain the geometry and pose of the user’s face for each frame by tracking the face using the ‘Faceshift’ software and a Prime sense Carmine 1.09 RGBD sensor. Calibration between the RGBD sensor and the color camera, which is used to capture the user’s input image sequence, is performed with a camera calibration toolbox.2 The camera’s intrinsic and extrinsic parameters let us align the face geometry with the input color images. Next, we manually position the eyeglasses onto the face mesh for the first frame. For all the examples we tested, this process took less than 5 minutes on average. A fully automatic option would use affine transformation computed based on preselected feature points on face and eyeglasses 3D model3 or a physics-driven technique.
After the initial manual positioning of the eyeglasses for the first frame, we track the head pose to automatically align the eyeglasses with the user’s face in the subsequent frames. This is achieved by calculating the relative pose change in each frame.

Parametric Lens Model:
Given the user’s eyeglasses prescription, we generate the 3D lens geometry based on a parametric model so that the optical power of the virtual lens corresponds to the user’s prescription. A lens is a

3D transparent and closed object. It consists of two main surfaces: a front surface and a back surface. The lens thickness is defined as the distance between front and back surface along its optical axis. Physical lenses are made of a transparent material with a certain refraction index, which affects lens thickness, weight, and optical properties.

Video Synthesis

In this stage the virtual eyeglasses are inserted into the input sequence using image-based rendering, where the eyeglasses are rendered using ray tracing. From the previous stage, we obtain a parametric lens model, the well-positioned eyeglasses, and the user's face geometry for each image frame. We address the rendering process by first describing the objects in the virtual scene and the materials associated with them. Then, we describe the ray-tracing rendering of the lenses, with the refraction and reflection effects and the shading cast on the user's face.

Conclusion:

Corrective lenses introduce distortion caused by the refraction effect, which changes the wearer's appearance. To give users a more realistic experience, a virtual try-on system modifies an input video and virtually inserts prescription eyeglasses with reflections and shading, producing an output similar to a virtual mirror. The authors also referenced a secondary paper that developed a novel algorithm that does robust real time tracking that is achieved by building suitable user specific blend shape models and exploiting the different characteristics of the acquired 2D image and 3D depth map data for registration.

The Authors of the primary paper solved a practical problem faced in everyday lives and it can easily be implemented in a large scale. The idea of virtual try on glasses with special considerations on the refraction properties of the lens with respect to the power of the lens has not been studied or implemented makes this a strong paper and a possible patent. This paper claims that the system is implemented have had successful user trails for 22 people with great success. So, in conclusion it's a great technology to be adapted into the computer visualizations for the future.

References:

Primary Paper:

Q. Zhang, Y. Guo, P. Y. Laffont, T. Martin and M. Gross, "A Virtual Try-On System for Prescription Eyeglasses," in IEEE Computer Graphics and Applications, vol. 37, no. 4, pp. 84-93, 2017.

doi: 10.1109/MCG.2017.3271458

Secondary Paper:

Weise, T., Bouaziz, S., Li, H., Pauly, M. 2011. Realtime Performance-Based Facial Animation. *ACM Trans. Graph. 30*, 4, Article 77 (July 2011), 9 pages. DOI = 10.1145/1964921.1964972 http://doi.acm.org/10.1145/1964921.1964972