COMPUTER GRAPHICS I
COMP.5460

LITERATURE REVIEW I

PRANITHA KEERTHI
01719717

LIGHTWEIGHT EYE CAPTURE USING A PARAMETRIC MODEL

This work is a parametric model of 3D eyes that is built out of a high resolution based database which scans both geometry wise and texture wise. This method provides both automatic reconstruction of real-eyes and artistic control over the parameters to generate user-specific eyes.

This work is mostly based on the work of Beard et al [2014] who developed a technique to scan the human eyes in high quality with the use of custom capture setup and a trio of algorithms which was tailored to handle the white sclera, the transparent cornea, and the deforming iris; the three visible parts of the eye. So, this work uses training database provided by Beard et al to build a parametric model of eye.

Aspects like the shape or the color of the eye can be controlled without in-depth knowledge of the subtleties of real eyes. The algorithms included in this paper touch on certain fields like, non-rigid alignment to find correspondence between eye meshes, data driven fitting to adjust an eye model to a given eye mesh, as well as constrained textured and geometry synthesis to create missing details not present in the acquired images. The algorithm of Li et al. [2008] has been used for the non-rigid alignment and the texture synthesis of the eyes.

As eye is composed of several components, a three separate component based model has been constructed, i.e., an eyeball mode (it represents the low-frequency variability of the entire eyeball shape), an iris model (it represents the high-resolution shape, color and pupillary deformation of the iris) , and a clear vein model (it represents the detailed vein structure in the sclera).

The eye database is given as the input data which contains eyes of different iris color ranging from brown to green-brown to blue, and the high resolution geometry captures intricate eye-specific surface details. The data provided contains also a limbus opacity mask defining the transparency transition from sclera to cornea, from which the position of the limbus can be extracted by mapping the 50 percent opacity level to the mesh.

To avoid overfitting to the samples, Principal Component Analysis(PCA) is used to reduce the dimensionality. PCA computes the mean shape plus a set of mutually orthogonal basis vectors from the samples, ordered according to their variance.

The algorithm iteratively refines the model in three steps, first by fitting the previous guess of the model to the sample shapes, second by deforming this fit outside of the model in order to more closely fit the samples, and third by recomputing the model from these fits.

Now the author shifts to the most salient component of the eye, i.e. the iris. A large variety of irises exist in the human population, and the most dominant are brown, blue and, green. Iris have a wide range of variation, so for this the author accounts for the variability by parametrizing the iris using a low-resolution control map. Using this low-resolution control map the iris needs to be synthesized to a high-resolution texture. For this purpose, the high-resolution texture from exemplar patches from the database are composed, following the well studied image quilting approach introduced by Effros and Freemman [2001].

To synthesize a specific patch in the output, the algorithm can thus consider sample patches at any angle with similar radii. The only drawback of the polar coordinate representation is that the synthesized texture must smoothly wrap across the left and right image boundaries.

The algorithm also provides extrapolations to other pupil dilations, allowing to control the iris virtually after reconstructing the eye. The iris geometries in the database are encoded in cylindrical coordinates(angle/radius/height), which renders them compatible to the domain used for texture synthesis. Encoding the geometry using differential coordinates is convenient. Since the geometry is encoded in cylindrical coordinates, we need to scale the angular dimensions(radians) to render it compatible with the radial and height dimensions, both being in mm units only.

The next component of the eye that the author shifts to is, the vein. Veins also travel under the surface at varying depths, and deeper veins appear thicker and softer, while veins at the surface appear more pronounced. The vein network is represented in a tree structure where the primary level veins constitutes of the large veins, the most salient structures when looking from afar, and then these veins branches off into smaller second, third and lower level veins. The position offsets of the vein are defined by the colored noise function C_{offset} in the range of pink (x = 1) and red (x = 2) noise. The veins on the sclera grow from the back of the eyeball to the front, and hence in this algorithm it is grown from bottom to top. The appearance of vein is influenced by many different factors, such as diameter, how shallow it grows, its oxygenation, and others. The first most important factor is depth and then later thickness. Vein growth is governed by a step size and a direction d at every point. The step size is attenuated during growth by decay factor. Here, the parametric models of the eye are concluded, i.e. the eyeball, iris and vein.

Now, the author talks about how all the above parameters can be estimated automatically and how the required iris control map is extracted from various sources. There are two use case scenarios on which the authors focus,

1. Multi-view fitting: It demonstrates how the proposed method may be used to complement existing photogrammetric face scanners to augment the facial geometry that is inaccurate for the eye itself with the high-quality eye reconstructions.
2. Single Image Fitting: It demonstrates how this method can be used to compute eye geometry and textures from single, uncalibrated input images.

To conclude, the this model allows to manipulate the captured data as it is full parametric, which includes changes like physiological effects or controlling the pupil size. The limitations of this method observed are single-image reconstructions is difficult to identify , the sclera vein synthesis does not guarantee to produce vein networks that match any partial veins that might be visible in the input images .

REAL-TIME 3D EYELIDS TRACKING FROM SEMANTIC EDGES

This work proposes a generative eyelid model which decomposes eyelid variation into two low-dimensional linear spaces which efficiently represent the shape and motion of eyelids.

It is usually difficult to model the shape and motion of local eye regions because they are small and their motions involve heavy occlusions in fold regions. PCA has been used by recent works to model the overall shape of the eye region from 22 scans but eye reconstructions requires efficient estimation of their motions.

First, the authors concentrate on the linear eyelid models. This system first proposes two linear models, with two sets of 3D mesh bases, for representing the eyelid shape and pose in low dimensional subspaces. To avoid the issue of appearance and disappearance of the folds and bulges of the eyelids, two linear rigs sharing identical bases are determined,

1. Shape Linear Rig: Based on certain characteristics of the eyelids,
   Position(relative positions of the eyes on the face), Contour Shape(eyes are round or flat, wide or narrow, upturned or downturned), Double-fold(distances between the fold and the upper eyelid on different parts, and the strength of the fold), and Bulge(shape and strength of bulge beneath the eye),
   an artist builds a set of eyelids to cover the characteristic variations. First, a basic 3D eyelid mesh with a neutral identity in a neutral, open pose, is built and then the basic mesh is modified to generate a set of new meshes.
2. Pose Linear Rig: The pose rig is built manually by generating a set of bases to cover all possible geometric changes caused by eyelid poses.

Next, the detection and identification of the four edges (fold edge, top edge, bottom edge, and bulge edge) for the eyelid motion capture is discussed. A naïve extension to identify the four edges is to train the four detectors, each of which is individually trained to detect only one of the four edges. The hostically-nested edge detection is modified to exploit the correlation among the four edges. First, the edge detection and identification problem is formulated as a multi-chaannel edge detection problem and then a unified energy metric is proposed to jointly consider the four edges together in the network training.

Next, the authors speak about training. The training set contains of 194 eye region images form 48 identities and the corresponding four-channel ground truth edge maps. The pose changes are caused by eyelid motions, as well as gaze motions, because eyelids change with eyeball movements. In the training, the network is fine-tuned from an initialization of the pre-trained VGG-16 net model, and a standard stochastic gradient descent algorithm in the training is used.

Next, the reconstruction of the 3D eyelids of a user with their linear eyelid models and the four edges identified on each recorded frame is discussed. The eyelid reconstruction method is integrated into a real-time face tracking system, which reconstructs 3D global poses, face identities, and expressions, as well as the eye-ball shapes and gaze directions, all in real time.

The limitations of this method are, the artist-designed bases do not cover all the variations in the eyelids, only a small dataset for training the eyelid edge detector is collected, and this system still depends on depth information in face tracking.

Finally, this paper can be concluded that almost a near-complete face reconstruction with realistic eye-regions can be constructed.

SIMILARITY BETWEEN THE TWO PAPERS

The primary and the secondary paper both concentrates on development of 3D structures based on the regions surrounding the eyes. According to both the papers state-of-the-art-real-time face tracking systems still lack the ability to realistically portray subtle details of various aspects of the face, particularly the region surrounding the eyes. To improve this situation the primary paper proposes a technique to reconstruct the 3D shape and motion of eyelids in real time and the secondary paper presents a new parametric model of 3D eyes built from a database of high-resolution scans with both geometry and texture.

REFERENCES

PRIMARY PAPER:

REAL-TIME 3D EYELIDS TRACKING FROM SEMANTIC EDGES

ACM Transactions on Graphics (TOG) - Proceedings of ACM SIGGRAPH Asia 2017

SECONDARY PAPER:

LIGHTWEIGHT EYE CAPTURE USING A PARAMETRIC MODEL

ACM Transactions on Graphics (TOG) - Proceedings of ACM SIGGRAPH 2016