

# Ultrafast Facial Tracker Using Generic Cameras with Applications in Intelligent Lifestyle

Yung-Hui Li<sup>1</sup>(✉), Yuan-Ting Hu<sup>1</sup>, Jethro Shen<sup>1</sup>, Mihai Preda<sup>2</sup>, Andrei Drexler<sup>2</sup>,  
Carmen Sosoiu<sup>2</sup>, Dragos Florin Stanculescu<sup>2</sup>, Paul Liu<sup>1</sup>, and Joe Ye<sup>1</sup>

<sup>1</sup> ULSee Inc., 3F., No. 28, Ln. 128, Jingye 1st Rd., Zhongshan Dist., Taipei, Taiwan  
{li,yuanhu,jethro,paul,joe}@ulsee.com

<sup>2</sup> Holotech Studios, Str Emanoil Porumbaru 77, 4<sup>th</sup> Floor, Sector 1, 011424 Bucharest, Romania  
{mihai, andrei, carmen.sosoiu, florin}@facereg.com

**Abstract.** The core of Human-Computer Interaction (HCI) is to analyze and understand the user's intension, which can be mostly manifested from the facial movement and expression of the user. Hence, the stage facial detection and tracking is extremely important in an user-friendly interface between human and computer. In ULSee, we developed an ultrafast markerless facial tracking system which is robust to variation in environmental lighting, pose and occlusion. It can be run at a speed of 10 ms/frame on an iPhone 6S system. With such accuracy and speed, it can be used to support many intelligent HCI applications. In this work, we envision an intelligent lifestyle in the future that can be built upon the basis of the ULSee's ultrafast markerless facial tracker, ranging from virtual reality, augmented reality, real-time facial recognition and driver drowsiness detection. We believe, that through the joint force between ULSee's world-class tracker and our clients, more user-awareness HCI application will be invented and a new lifestyle will arise.

**Keywords:** Markerless facial tracking · Real-time avatar animation · Facial recognition · Driver drowsiness detection

## 1 Introduction

The core of HCI (Human-Computer-Interaction) system lies in its ability to understand user's intention and give an appropriate reaction based on the current environmental condition and context. To understand user's intention, most systems rely on the vision-based solution, which is to process the images/videos captured from the camera, locate the user's head, and further analyze the user's face image. There are multiple user attributes that an intelligent HCI system can infer based on vision-based solution, for example: user's emotion, age, sex, and expression variation. Those user's attributes together with dialog context can further be used as input for an intelligent HCI system. Therefore, how to locate user's face with high accuracy and analyze it is one of the most important functions in a vision-based and face-centric HCI system.

Face tracking has received a lot of attentions in research communities in recent years [1–7]. Traditional face tracking is to locate the rectangular area where the face appears

in a given input image. However, in order to further analyze user’s attributes mentioned in the previous paragraph, detailed localization information about each local part of a face (e.g., eyes, nose and mouth) is also needed. Such detailed localization information can be trained and modeled by so-called “shape models”, for example, active shape models and its variants [8–16]. A shape model, if constructed with robust machine learning algorithms and trained with large amount of data, can be used to track local part of a face very accurately and very efficiently (in real time). In ULSee, we implemented an ultrafast and very robust face tracker which can run in real time, and it works with generic consumer-level camera (i.e., no special camera needed). Besides facial tracking, there is a 3D face model built inside our tracker which can estimate the pitch, yaw and roll angle of the input face image in real time. The system flowchart of ULSee’s facial tracker is described in Fig. 1.

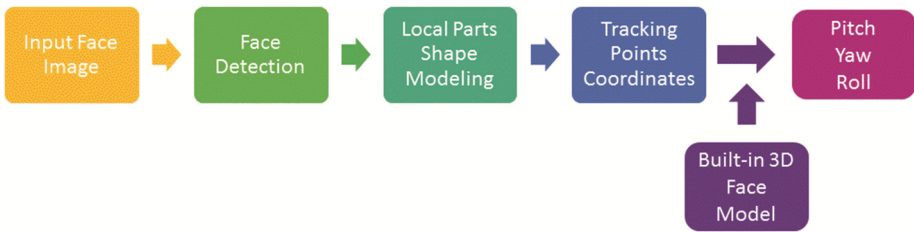


Fig. 1. The system flowchart of ULSee’s facial tracker.

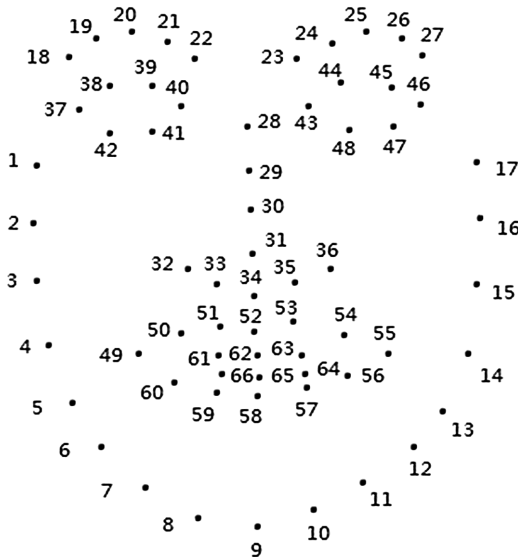


Fig. 2. The 66 tracking points for a face defined by ULSee facial tracker.

The number of tracking points returned by ULSee's tracker is 66. Those 66 points represent the location of lines that define the face contour (jaw line), eyebrows, eyes, nose and mouth accordingly. The numbering of these 66 points is illustrated in Fig. 2.

The output of ULSee's facial tracker (the coordinates of the 66 points, as well as the pitch, yaw and roll angles) can be further fed into other intelligent system for different applications. So far, ULSee's face tracker has been integrated with clients' system and many interesting and useful applications have been created, ranging from markerless real-time avatar animation, augmented reality (in the form of virtual glasses and jewelry try-on), face recognition and driver drowsiness detection. In the following sections, we will illustrate these applications one by one.

## 2 Real-Time Avatar Animation

One of the important clients to ULSee is Holotech Studios, the creator of FaceRig, which is a software that allows anyone to embody and animate outstanding real time CG character portraits via motion capture from a webcam. In FaceRig there are two distinct ways of translating tracking data to a 3D model movement.

### 2.1 Two Methods for Avatar Animation

One way is by interpreting tracking data as presence of certain landmark configurations which in turn should underlie actual human expressions, or more exactly human face postures. In this method, certain tracking data represents, for instance, a lifted eyebrow, or an eye squint, because respective landmarks change their relation in a specific way. For each distinct identifiable landmark configuration, a specific meaning can be attributed and this interpretation can be passed on to a system that establish correlations between these configurations and 3D animations states. We refer to this method as "animation retargeting" (or animation atomics retargeting).

Another way of transferring tracking data to a 3D model is by translating landmark movement to 3D object movement. Because the landmarks evolve in 2D space and because their inherent spatial relation is determined by the human face image on which those landmarks were identified, certain corrections and approximations must be done in order to function on a 3D model that not only moves in an extra dimension (depth) but very likely has different proportions between its inner features that the tracked human face. In this method the tracking data is not interpreted and thus does not use corresponding animations, but rather amplifies or diminishes position deltas in order to produce similar spatial relations in the 3D model components. In addition to these movement modifications, the depth placement and movement is approximated. This way the landmark movement is retargeted on the 3D model. We call this "free-form retargeting".

## 2.2 Animation Retargeting

Animation based retargeting, doesn't set 3D transformations (position, orientation, scale) directly to 3D objects. Tracking data is really a signal that certain landmark configurations are present. The configurations are relative to a neutral reference state and thus can be expressed in a normalized way, with their presence (actualization) having values between 0 (none) and 1 (maximum certainty that landmarks are different in a specific way). For each trackable feature a correspondent 3D spatial relation between 3D objects is made, carrying the same meaning, but the spatial relation between landmarks and 3D objects is arbitrary. They don't even have to correspond one to one. What mirrors the trackable features are 3D configurations not 3D objects. If the inner left eyebrow being raised is identifiable as such by the tracker, then there is a 3D configuration that mirrors this state, and the actualization of this expression state determines the actualization of the corresponding 3D configuration.

This 3D configurations, or spatial relations, represent the maximum tracking value. Because each trackable feature can have values between 0 and 1, and because the values are expressed relative to a tracking reference, the way correspondent 3D configurations behave must mirror this characteristic. They have to be expressed as an offset and this offset must be weighted, also between 0 and 1, just as the tracking feature.

In 3D graphics, the data storing attributes that vary their values over time are called animations. Thus, the additive 3D configurations with variable actualization are named additive animations. Neutral reference states are called base animations, because they serve as a 3D matrix base for final 3D space representation. While this neutral states could be static, and thus not really animations, it's probable that this neutral states to augment what is being tracked by additional movement, becoming actual animations.

The base animations are tracking independent, while the additive animations simply represent the trackable features in 3D space. Because the 3D models don't usually consist just in a collection of discrete objects, but rather mimic the appearance of real or realistic objects, the necessity of another approximation appears. The translation is from animated 3D objects and rendered 3D objects. The animated 3D objects serve as a base of translating tracking data in 3D data. 3D render objects depict real or realistic objects and they get to be drawn by the 3D render.

FaceRig tracking data analysis is based on the Facial Action Coding System (FACS) developed by anatomist Carl-Herman Hjortsjö and later updated by Ekman, Friesen and Joseph C. Hager. Facial movements are encoded by FACS in basic actions of individual muscles or groups of muscles called Action Units (AU). Some of the Action Units recognized by FaceRig are the following:

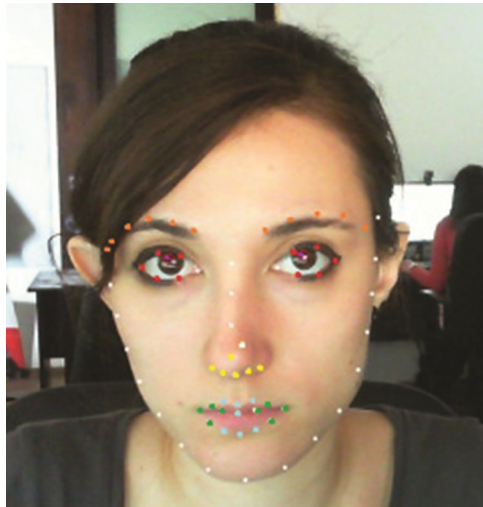
- Inner Brow Raiser
- Outer Brow Raiser
- Lid Tightener
- Nose Wrinkler
- Upper Lip Raiser
- Lip Corner Depressor
- Lip Pucker
- Jaw Drop

- Mouth Stretch
- Eyes Closed
- Head Turn Left/Right
- Head Up/Down
- Head Tilt Left/Right

For each Action Unit weight or intensity of the corresponding facial movement is computed. The weight is generally computed as a deviation of a specific group of landmarks' transformations from a default neutral pose.

In the Fig. 3 we are illustrating groups of landmarks that contribute to Action Units computation as follow:

- The red group of landmarks contribute on Lid Tightener and Eyes Closed AUs
- The orange group of landmarks contribute on Inner Brow Raiser and Outer Brow Raiser AUs
- The yellow group of landmarks contribute on Nose Wrinkler and Upper Lip Raiser AUs
- The blue group of landmarks contribute on Nose Wrinkler, Upper Lip Raiser, Jaw Drop, Lip Corner Depressor and Lip Pucker AUs
- The green group of landmarks contribute on Mouth Stretch, Lip Pucker, Lip Corner Depressor and Mouth Stretch AUs



**Fig. 3.** The correspondent relation between the groups of landmarks and Action Units (AU) (Color figure online)

The next step in the Animation Atomics Method is to map each Action Unit weight to a keyframe of the corresponding atomic animation of the avatar. After each animation is set to the correct keyframe they are blended together with the help of an Animation Tree.

### 2.3 Free-Form Retargeting

Free-form retargeting, at its core, is a system that translates image-space landmark movement (relative to a user-calibrated neutral pose) to 3D bone movement (relative to an artist-defined neutral pose). In its current incarnation in the FaceRig application, free-form retargeting can drive the bones corresponding to the avatar's mouth and eyebrows directly from tracking information. For aesthetic reasons a set of secondary bones, for which no tracking information is available, can also be driven by analyzing the current pose as defined by the primary bones and choosing the best matches from a set of artist-defined frames, enhancing expressivity (e.g. by simulating the thickening and rounding off of the cheeks when smiling). Since free-form retargeting only affects a few specific areas and not the entire bone hierarchy of the avatar it is only used in addition to Animation Atomics, overriding its output for the controlled regions.

The system relies on naming conventions for bone identification. For the mouth region a supplementary UV-mapped support mesh is used to define the area of movement for the affected bones. Secondary bone movement requires special animations, where each frame is interpreted as a specific pose to be identified and applied dynamically.

During avatar initialization, the neutral matrix in the idle pose for each controlled bone is recorded. For the mouth bones, the corresponding 2D location in the UV-space of the support mesh is also computed and stored. The 3 directly controlled regions (mouth, left and right eyebrows) are measured, allowing the system to drive avatars with different proportions from the user's.

Two examples of real-time avatar animation using FaceRig and ULSee's tracker are given in Fig. 4.



**Fig. 4.** Example images of real-time avatar animation using FaceRig with ULSee's facial tracker.

## 3 Virtual Try-on

### 3.1 Goal

Virtual try-on has attracted great attentions in recent years because of its commercial value in on-line shopping. A virtual try-on system with 2D cameras is much more favorable than using 3D sensors due to the availability of 2D cameras. However, estimating the 3D geometry via 2D images has many ambiguities and the inserted virtual objects are hardly fit to the real scenes, which degrades the user experience of the virtual

try-on system. To this end, we develop a robust real-time virtual try-on system, the ULSee VTO system, based on our accurate facial tracker using only 2D consumer cameras. Our virtual try-on system consists of virtual glasses try-on, virtual cosmetics try-on, virtual jewelry try-on and other related face/head virtual try-on. The ULSee VTO system provides a practical 2D solution for virtual try-on and demonstrates its superior effects on the aforementioned try-on tasks.

### 3.2 Literature Review

Methods for real-time virtual try-on can be separated into two categories with respect to the two different image capturing devices, namely the 2D RGB cameras [18, 19, 21] and the 3D RGBD sensors [17, 20, 22]. Using the 3D sensors can get more accurate estimation than using the 2D cameras as it contains depth information, while the availability of the 3D sensors and its fragility to sunlight makes the 3D virtual try-on system less accessible. Despite the difference of the type of sensors, both 2D and 3D solutions rely on accurate tracking to capture the motion in real scene and reflect to the user in real-time. A fast and accurate facial tracker is an essential component in a virtual try-on application because it provides accurate head pose and landmark location estimation.

### 3.3 Tracking for Virtual Try-on

The ULSee VTO system contains virtual glasses try-on, virtual cosmetics try-on, virtual jewelry try-on and other related face/head virtual try-on. The virtual try-on system uses RGB images as input so that the system can run on devices such as mobiles and pads.



**Fig. 5.** The results of 6 different pairs of glasses try-on in the ULSee VTO system.

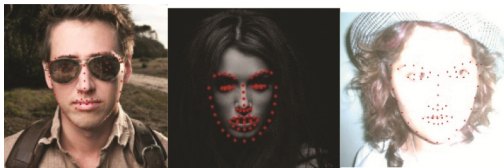
At the core of the ULSee VTO system is the ULSee facial tracker and the system uses the head tracker and head pose estimation to transform the virtual objects to fit the user's pose. To avoid jitters of the inserted objects, a further smooth pose will be calculated based on the temporal information. Then virtual objects, such as glasses, jewelries and masks, can be placed into the 2D image with respect to the landmark location given by the tracker.

The ULSee VTO system provides a much more convenient solution to reduce the efforts of try-on in physical stores. The examples of virtual glasses try-on using our system are shown in Fig. 5. The realistic try-on results demonstrate the effectiveness of the ULSee tracker.

## 4 Face Recognition

### 4.1 Difficulties in Face Recognition

In facial analysis, including face recognition, emotion recognition, and face demographics, three key issues need to be considered. The first is the pose of the face. If the faces are not frontal, no matter what methods are used, the face recognition rate will be seriously degraded. This is called pose alignment issue. The second issue is how to align gallery and probe face images correctly. For any given face, a few anchor points can always be retrieved. If the anchor points of each face image did not align to some specific positions, the feature extracted in local region will not match to each other, which in turn, degrades the recognition rate. This is called 2D alignment issue. The third issue is the occlusion due to sunglasses, eyeglass, and scarf. In such cases, the image features of the occluded areas can not be retrieved, which impact recognition rate as well. Using ULSee's facial tracker, we will be able to fix all above issues. As shown in Fig. 6, ULSee's facial tracker is able to predict the position of each anchor point very accurately and estimate face pose precisely to help pose alignment. For the case of occlusion, because ULSee's tracker can return confidence scores for each anchor points, the confidence scores can be used as a weighted coefficient if recognition-by-parts algorithms are used.



**Fig. 6.** ULSee's face tracker can estimate precise facial points locations in dark, strong lighting condition, and wear sunglasses. Above images are acquired from internet.

### 4.2 Related Works

There are 2D and 3D approaches for face alignment. In [23, 24], the researchers showed that the effective alignment could improve recognition rate. For 2D alignment methods, the



objective focus on aligning several anchor points of each face image to some specific positions. This process will enhance the discriminability of face image. In 3D alignment methods, the general objective focuses on changing non-frontal face images to frontal face images. In this transformation, it needs accurate estimations for the angle of face pose. If the estimated face pose is not accurate enough, the transferred face image may be broken.

### 4.3 ULSee's Approach

In ULSee, our face recognition system can achieve 97.41 % recognition rate in Labeled Faces in the Wild (LFW) view2 dataset [26] by using the advantages of our tracker. The advantages lie in that the tracker can detect and track 66 fiducial points precisely in real-time with estimated confident values of 66 points. For example, if users wear sunglasses, the confident values of points on eye region may drop to 0.3. When using recognition-by-parts method with these confident values we can get better recognition performance.

## 5 Driver Drowsiness Detection

### 5.1 Goal

When driving a car, it is very easy for the driver to be distracted or get tired. Sensor-based technology has been developed specifically to help the driver avoid such situation, as shown in Fig. 7. This kind of technique needs constantly analyze the driver's face pose to tell if it is frontal or not. If the driver does not look straight ahead or her eyes are constantly closed up, it is highly likely that she is falling asleep [25]. This mechanism minimizes the risk due to driver's distraction or fatigue. Such application needs robust facial tracking and accurate pose estimation under different environmental illuminations, such as extremely weak or strong lighting condition.



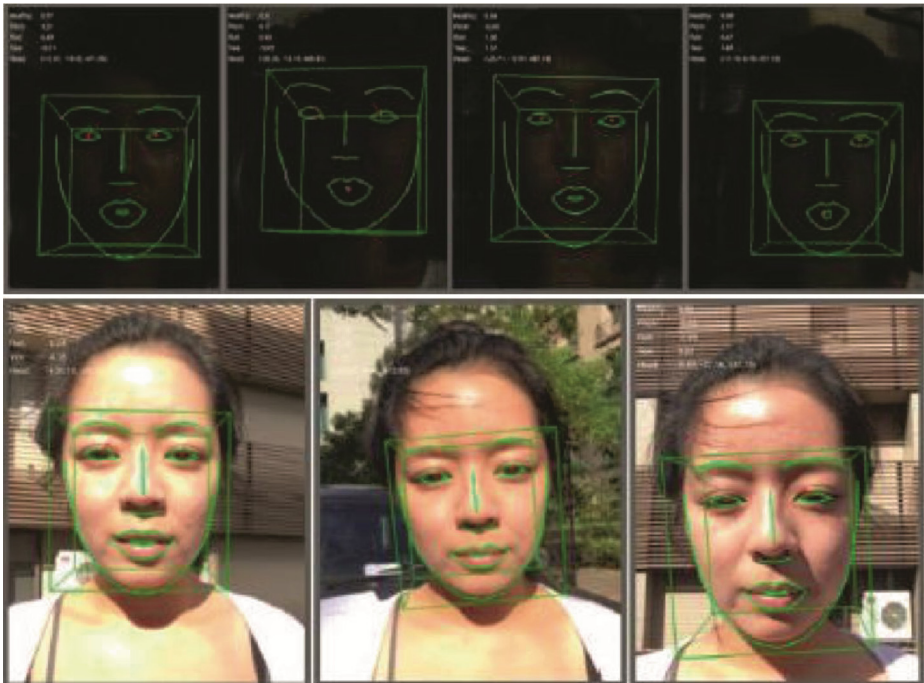
**Fig. 7.** The illustration of the concept of driver drowsiness detection system, which is borrowed from [25].

### 5.2 Related Works

There are many facial tracking techniques in the literature. Some methods focus on fast point's detection. Some methods focus on fast point's tracking. Most of them can not handle the problem due to different lighting conditions.

### 5.3 ULSee's Approach

In ULSee, our facial tracker can deal with various illumination variations and estimate the head pose in real-time (more than 30 fps). Our tracker is able to operate in a wide range of environmental illumination, from 0.02 to 60,000 lx, as shown in Fig. 8. Besides the pose, our tracker is also able to return the precise location for eyes and eyebrow. Such robustness enables our tracker to be the core of the technology for the driver drowsiness detection. The final system is able to send out loud alarm when the driver seems to doze (through eyes contour analysis) or distracted (by analyzing the head pose). It is currently under internal testing with our client, who is one of the major automakers in Europe.



**Fig. 8.** The facial tracking result under extremely lighting variation (provided by ULSee). Top row: tracking result taken at 0.02 lx. Bottom row: tracking result taken at 46,000-60,000 lx.

## 6 Conclusion

Imagine a new future era driven by computer vision technology. When you drive, the intelligent system in car automatically monitors your status and alert you when you are

distracted or drowsy. When you get home, the surveillance system automatically recognizes you are the owner of the house and open the door. When you get inside the living room, the air conditioner automatically turns on and the temperature is set to what you usually want. The TV is also turned on and tuned to your favorite sport channel. After dinner, when you play an online game with your friend, you can choose Yoda as your avatar, and Yoda will demonstrate exactly the same facial expression as you squint your eye or laugh widely.

With an ultrafast and robust facial tracking technology (for example, ULSee's facial tracker), all of the above applications are on their way and will be happening in the near future. User-friendly applications will make the future home and office become user-aware, and user-oriented intelligent lifestyle will become a mainstream.

## References

1. Saxena, V.; Grover, S.; Joshi, S.: A real time face tracking system using rank deficient face detection and motion estimation. In: 7<sup>th</sup> IEEE International Conference on Cybernetic Intelligent Systems, 2008 (CIS 2008), pp. 1–6, 9–10 September 2008. doi:[10.1109/UKRICIS.2008.4798956](https://doi.org/10.1109/UKRICIS.2008.4798956)
2. Harguess, J.; Changbo Hu; Aggarwal, J.K.: Occlusion robust multi-camera face tracking. In: 2011 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 31–38, 20–25 June 2011. doi:[10.1109/CVPRW.2011.5981790](https://doi.org/10.1109/CVPRW.2011.5981790)
3. Yoder, J., Medeiros, H., Park, J., Kak, A.C.: Cluster-Based Distributed Face Tracking in Camera Networks. *IEEE Trans. Image Process.* **19**(10), 2551–2563 (2010). doi:[10.1109/TIP.2010.2049179](https://doi.org/10.1109/TIP.2010.2049179)
4. Faux, F., Luthon, F.: Robust face tracking using colour Dempster-Shafer fusion and particle filter. In: 2006 9th International Conference on Information Fusion, pp. 1–7, 10–13 July 2006. doi:[10.1109/ICIF.2006.301713](https://doi.org/10.1109/ICIF.2006.301713)
5. Wang, P., Ji, Q.: Robust face tracking via collaboration of generic and specific models. *IEEE Trans. Image Process.* **17**(7), 1189–1199 (2008). doi:[10.1109/TIP.2008.924287](https://doi.org/10.1109/TIP.2008.924287)
6. Painkras, E., Charoensak, C.: FaceProcessor: a framework for hardware design and implementation of a dynamic face tracking system. In: 2005 Fifth International Conference on Information, Communications and Signal Processing, pp. 172–176 (2005). doi:[10.1109/ICICS.2005.1689028](https://doi.org/10.1109/ICICS.2005.1689028)
7. Shi, L., Zhu, Y.: Robust face tracking-by-detection via sparse representation. In: 2015 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), pp. 1–4, 19–22 September 2015. doi:[10.1109/ICSPCC.2015.7338846](https://doi.org/10.1109/ICSPCC.2015.7338846)
8. Cootes, T.F., Taylor, C.J.: Using grey-level models to improve active shape model search. In: Proceedings of the 12th IAPR International Conference on Pattern Recognition, 1994. Vol. 1 - Conference A: Computer Vision & Image Processing, pp. 63–67, 9–13 October 1994. doi:[10.1109/ICPR.1994.576227](https://doi.org/10.1109/ICPR.1994.576227)
9. Sugawara, Y., Lee, D.S., Kawanaka, A.: 3-D shape model retrieval using multi range image phase correlation method. In: 2012 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), pp. 000061–000066, 12–15 December 2012. doi:[10.1109/ISSPIT.2012.6621261](https://doi.org/10.1109/ISSPIT.2012.6621261)
10. Davies, R.H., Twining, C.J., Cootes, T.F., Taylor, C.J.: Building 3-D statistical shape models by direct optimization. *IEEE Trans. Med. Imaging* **29**(4), 961–981 (2010). doi:[10.1109/TMI.2009.2035048](https://doi.org/10.1109/TMI.2009.2035048)

11. Luo, S., Li: Accurate object segmentation using novel active shape and appearance models based on support vector machine learning. In: 2014 International Conference on Audio, Language and Image Processing (ICALIP), pp. 347–351, 7–9 July 2014. doi:[10.1109/ICALIP.2014.7009813](https://doi.org/10.1109/ICALIP.2014.7009813)
12. Cootes, T.F., Taylor, C.J., Lanitis, A.: Multi-resolution search with active shape models. In: Proceedings of the 12th IAPR International Conference on Pattern Recognition, 1994. Vol. 1 - Conference A: Computer Vision & Image Processing, pp. 610–612, 9–13 October 1994. doi:[10.1109/ICPR.1994.576375](https://doi.org/10.1109/ICPR.1994.576375)
13. Baloch, S.H., Krim, H.: Flexible Skew-Symmetric Shape Model for Shape Representation, Classification, and Sampling. *IEEE Trans. Image Process.* **16**(2), 317–328 (2007). doi:[10.1109/TIP.2006.888348](https://doi.org/10.1109/TIP.2006.888348)
14. Neumann, A.: Graphical Gaussian shape models and their application to image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **25**(3), 316–329 (2003). doi:[10.1109/TPAMI.2003.1182095](https://doi.org/10.1109/TPAMI.2003.1182095)
15. Huang, H., Makedon, F., McColl, R.: High dimensional statistical shape model for medical image analysis. In: 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008, pp. 1541–1544, 14–17 May 2008. doi:[10.1109/ISBI.2008.4541303](https://doi.org/10.1109/ISBI.2008.4541303)
16. Igual, L., De la Torre, F.: Continuous procrustes analysis to learn 2D shape models from 3D objects. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 17–22, 13–18 June 2010. doi:[10.1109/CVPRW.2010.5543280](https://doi.org/10.1109/CVPRW.2010.5543280)
17. Giovanni, S., Choi, Y.C., Huang, J., Khoo, E.T., Yin, K.: Virtual try-on using kinect and HD camera. In: Kallmann, M., Bekris, K. (eds.) MIG 2012. LNCS, vol. 7660, pp. 55–65. Springer, Heidelberg (2012)
18. Huang, S.H., Yang, Y.I., Chu, C.H.: Human-centric design personalization of 3D glasses frame in markerless augmented reality. *Adv. Eng. Inf.* **26**(1), 35–45 (2012)
19. Saragih, J.M., Lucey, S., Cohn, J.F.: Real-time avatar animation from a single image. In: IEEE International Conference on Automatic Face and Gesture Recognition and Workshops (2011)
20. Tang, D., Zhang, J., Tang, K., Xu, L., Fang, L.: Making 3D eyeglasses try-on practical. In: IEEE International Conference on Multimedia and Expo Workshops (2014)
21. Yuan, M., Khan, I.R., Farbiz, F., Niswar, A., Huang, Z.: A mixed reality system for virtual glasses try-on. In: Proceedings of the 10th International Conference on Virtual Reality Continuum and Its Applications in Industry (2011)
22. Zhu, X., Qin, S., Yu, H., Ge, S., Yang, Y., Jiang, Y.: Interactive virtual try-on based on real-time motion capture. In: Advances in Multimedia Information Processing (2012)
23. Chen, D., Cao, X., Wen, F., Sun, J.: Blessing of dimensionality: high-dimensional feature and its efficient compression for face verification. In: Computer Vision and Pattern Recognition (CVPR) (2013)
24. Hassner, T., Harel, S., Paz, E., Enbar, R.: Effective face frontalization in unconstrained images. In: Computer Vision and Pattern Recognition (CVPR) (2015)
25. <http://www.paneuropeannetworks.com/special-reports/eyealert-driver-fatigue-detection/>
26. Huang, G.B., Learned-Miller, E.: Labeled Faces in the Wild: Updates and New Reporting Procedures. Technical Report UM-CS-2014-003, University of Massachusetts, Amherst, May 2014