Face and Eye-ware Classification using Geometric Features for a Data-driven Eye-ware Recommendation System

Amir Zafar*
Concordia University

TIberiu Popa†
Concordia University

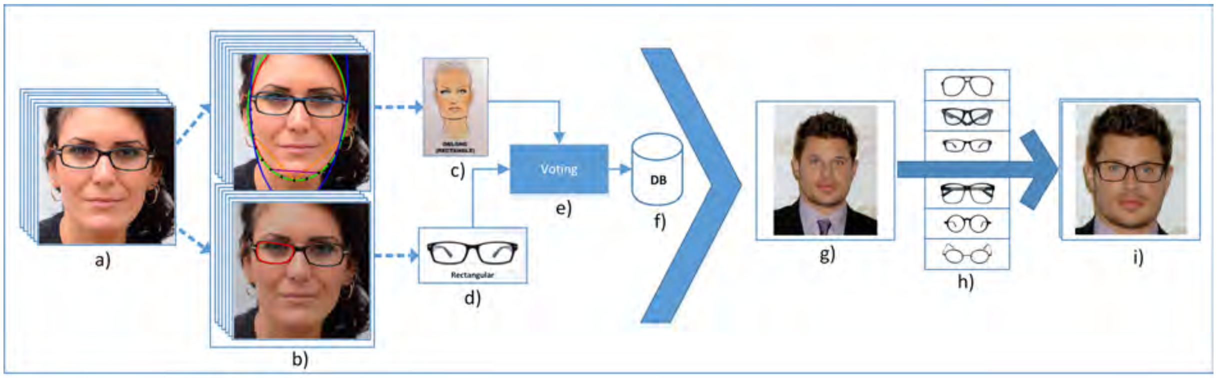

Figure 1: Our eye-ware recommendation system. a-f) Data collection pipeline. g-i) Recommendation pipeline. In detail: a) An image database of persons wearing eye-ware. b) Top: Geometric features extracted from the face. Bottom: Automatic eye-ware shape extraction. c) Classification of face shape. d) Classification of eye-ware shape. e-f) Rating of images from a survey are stored together with the face and eye-ware classification. We extract the most popular combinations. g) An input image of a person without eye-ware. h) Face shape detection is performed on the image and based on the results from (f) a recommendation is made. i) A eye-ware frame of the recommended type is overlapped to the face.

Abstract

In this work we present an automatic shape extraction and classification method for face and eye-ware shapes. Our novel eye-ware shape extraction algorithm can extract the polygonal shape of eye-ware accurately and reliably even for reflective sun-glasses and thin metal frames. Additionally, we identify key geometric features that can differentiate reliably the shape classes and we integrate them into a supervised learning technique for face and eye-ware shape classification. Finally, we incorporate the shape extraction and classification algorithms into a practical data-driven eye-ware recommendation system that we validate empirically with a user study.

Index Terms: I.3.8 [Computer Graphics]: Applications;

1 Introduction

Despite the gain in popularity of laser surgery and contact lenses, around 64% of the population still wears corrective eye-glasses, half of them wears them all the time. Additionally, sun-glasses are now ubiquitous accessories, therefore it is fair to say that eye-ware, including both corrective eye-glasses and protective sun-glasses, are likely the most popular facial accessory if both genders are considered. As a result, nowadays eye-ware has become a fashion item with iconic designs and an overwhelming supply of frames is available in both online and tradition retail stores, opening the path to eye-ware recommendation systems [1 2]. Most current recommendation systems, however, are generally procedural, based on some a priori aesthetic rules and not necessarily on real user data. These aesthetic rules are opaque to the user, occasionally contradicting each other, and are based on rigid aesthetic principles that do not take into account current fashion, age or culture.

In this work we present a scalable and fully automatic data-driven eye-ware recommendation system that employs a database of user tagged images in order to extract the most popular combinations of face and eye-ware shapes. Our recommendation system has two phases as illustrated in Figure 1: a training phase (Figure1(left)) and a testing phase (Figure1(right)). In the training phase, a database of images featuring persons wearing eye-ware is created. Next, the face and eye-ware shapes of the persons are automatically extracted and classified into a set of standard shapes as illustrated in Figures 2 and 3. This classification system seems to be ubiquitous in the fashion industry and it encompasses most existing shapes. A survey is then collected where a set of persons are asked to rate the aesthetic compatibility between the face and eye-ware shapes. In the testing phase, the input is an image of a person without wearing eye-ware. The face shape is automatically extracted and a recommendation of one or two frames is made based on the most popular combination obtained from the survey. We validated our recommendation results in a follow-up study where we used an online virtual eye-ware try-on system for visual feedback to recommend eye-ware. Our contributions in this work are:

- an automatic face classification system based on a single image, using a robust tailored set of geometric features;
- an automatic polygonal extraction of the eye-ware shape from a single image;
- an automatic frame classification system based on a single image, using a robust tailored set of geometric features; and
- an automatic data driven eye-ware recommendation system validated by a user study.

*e-mail: amirzafarasoodeh@gmail.com
†e-mail:tiberiu.popa@concordia.ca

[1]`http://www.lenscrafters.com/lc-us/face-shape`
[2]`http://glassescrafter.com/`

Graphics Interface Conference 2016
1-3 June, Victoria, British Columbia, Canada

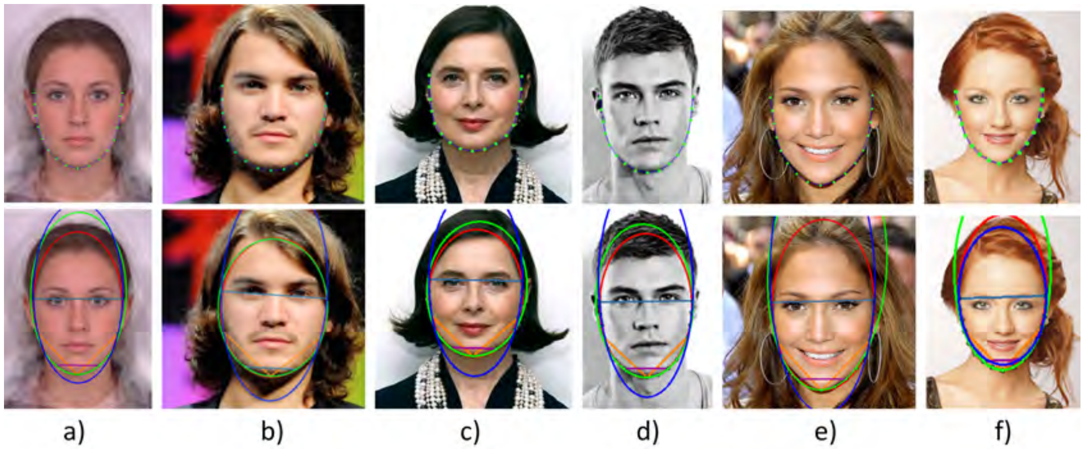

Figure 2: Illustration on how the geometric features vary with the face type: a) oval, b) round, c) square, d) oblong, e) diamond, f) heart. Top: Images with face boundaries outlined. Bottom: Images with the extracted geometric features. Red: Ellipse that best fits all boundary points. Green: Ellipse that best fits the chin points. Blue: Ellipse that best fits the cheek boundary points. Orange: Diagonal line. Purple: Jaw line. Lighter blue: Eye line.

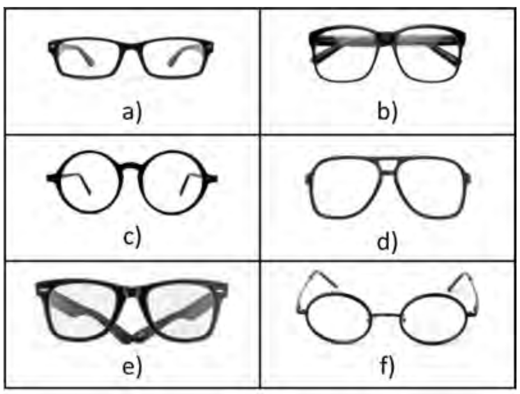

Figure 3: Example of the six standard frame shapes: a) rectangular, b) square, c) round, d) aviator, e) wayfarer, f) oval.

2 Related work

The perceptual aspect of facial attractiveness has been studied thoroughly by social psychologists [5] and has many applications in medical fields such as plastic surgery [10] and in the fashion industry to determine the most attractive hairstyle [3], makeup or eye-ware. In tandem with social and medical disciplines, computational disciplines tried to quantify and automate the detection of salient facial features. For instance, some work focuses on quantifying the attractiveness of a person automatically based on a single image [9, 3]. Furthermore, applications in face detection, face feature extraction and authentication based on a single image have also received a lot of attention [13, 14]. In this work the main focus is on the aesthetic correlation between the shape of the face and the shape of eye-ware. We employ a data-driven approach that has at its core a robust and fully automatic face and eye-ware shape extraction and classification methods. We look at these components one by one.

[3] `http://www.allure.com/`

2.1 Face shape extraction and classification

While recent state of the art face extractors [13] are very reliable to automatically detect from a single image the important features such as face boundary and salient points such as the eyes, face shape classification still remains a difficult problem. Many general shape classification methods have been proposed [16, 15]. Depending on the flavour, they work well if the shapes are significantly different to each other. In our case, however, the shape variation from one category to another is often very subtle and some face shapes cannot be classified as one shape, but rather as a blend between two shapes.

Therefore, a tailored approach is required in order to differentiate the subtle characteristics of each shape category and to account for the blending of the shapes. Such a variation is illustrated in Figure 2.

2.2 Eye-ware shape extraction

The first practical shape extraction method for eye-ware was introduced by [7]. Their method is not very reliable with only about 50% accuracy. Several methods are targeted at eye-ware detection [17, 8] or eye-ware removal from the image [11, 6]. These methods typically do not extract an accurate outline of the eye-ware shape and thus are not suitable for our application. A purely geometric method that requires no a priori knowledge was developed by [18], based on Delauney triangulation of the edges in the image. The results usually contain noise and outliers, which can affect the face classification. More recently, Borza et al. [2] presented a reconstruction method based on template matching of a database of pre-existing shapes. This is inadequate for our application as any new eye-ware designs will have to constantly be introduced to the database. The eye-ware classification problem suffers from the same challenges as the face classification: shape changes between different classes are subtle, making it difficult to differentiate using a generic shape classification algorithm. A more accurate and robust method tailored for this specific problem is desirable. We chose a case based reasoning (CBR) [1, 12] framework approach that is suitable for both the face shape classification as well as the eye-ware shape classification.

The rest of the paper is organised as follows: Section 3 presents the face classification method. Section 4 presents the eye-ware shape extraction method. Section 5 presents the eye-ware classification method. Section 6 presents the data-driven eye-ware recommendation system. Section 7 presents the discussion and Section 8

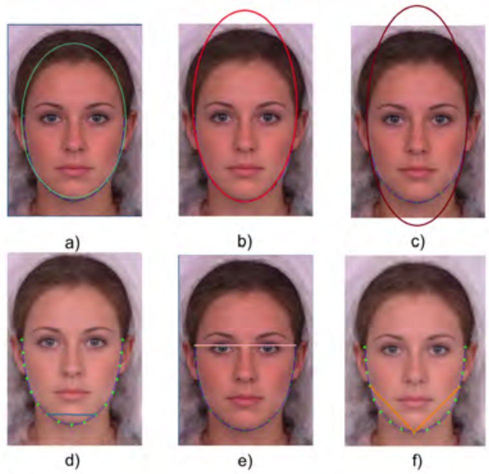

Figure 4: Geometric features used for the face shape classifier. a) Best fitting ellipse to all the face boundary points (G1). b) Best fitting ellipse to the chin boundary points (G2). c) Best fitting ellipse to the cheek boundary points (G3). d) Jaw line (G4). e) Eye line (G5). f) Diagonal lines (G6).

presents the conclusions, limitations and future work.

3 Face shape classification

The face shape classification consists of four steps. First, using a state of the art face tracker [13], the face boundary polygon is extracted from the image. A set of geometric features tailored to differentiate between the six face shape types are identified and computed from the face boundary (Figure 4). These geometric features are converted into a feature vector that can be used computationally in the CBR framework. Using this feature vector, the CBR framework is trained on a set of known face shapes to learn a distance metric defined on the feature vector. Finally, the classification is done based on nearest neighbour search in the training set using the learned distance.

3.1 Geometric features extraction

The face classification into five classes: oval, round, square, oblong (rectangular), diamond and heart, depends only on the boundary polygon points extracted from the face tracker. These classes are widely used in the fashion industry. Based on the characterisation and study of each face shape, we identified experimentally a set of key geometric properties illustrated in Figure 4: best fitting ellipse to all the face boundary points (G1), best fitting ellipse to the chin boundary points (G2), best fitting ellipse to the cheek boundary points (G3), jaw line (G4), eye line (G5) and diagonal lines (G6). The choice of these features is rooted in the observed geometric relationship between the six shapes. (G1) helps identify the oval face measuring how well the boundary points fit the ellipse. (G2) helps identify faces with pointy chin such as diamond and heart. For these face types, the height of this ellipse will be relatively large (Figure 2e,f). (G3) helps identify square and oblong faces where the height of this ellipse will be relatively large (Figure 2c,d). (G4) and (G6) also help identify faces with pointy chins such as diamond and heart, while (G5) is used for normalisation. Our algorithm is not very sensitive to the computation of these features, therefore a rough approximation is sufficient. The most difficult classes to differentiate are between diamond and heart. In our experiments (G3) can discriminate between them.

Based on the observations mentioned above, we incorporate these geometric features into our CBR framework by creating an eight dimensional feature vector. The first three components of the feature vector ($F_1 - F_3$) are the height of the three ellipses (G1, G2, G3) in Figure 4a-c) normalised by the eye-line (G5). The second set of three components of the feature vector ($F_4 - F_6$) are the sum of the distances of the face boundary to the three ellipses (G1, G2, G3) normalised by the eye-line (G5). The last two features are the average jaw line (F_7) (Figure 4d) and the average diagonal line (F_8) (Figure 4f) both normalised by the eye line (Figure 4e). The components of the feature vector were determined experimentally on a set of given data.

3.2 Training

This eight dimensional feature vector is computed on a training set of images whose face shapes are known. The remaining ingredient is the distance metric. As the relative contribution of each feature to the distance function is unknown, we define the distance function as follows:

$$d(T^1, T^2) = \sum w_i (T_i^1 - T_i^2)^2 \tag{1}$$

where T^1, T^2 are two feature vectors extracted from an image, the subscript i denotes the components of the eight dimensional feature vector and w_i is a weight associated to each of the features to compensate for the relative magnitude and importance of the feature. We compute the weights iteratively using the following process. We set all weights to 1. We introduce each training sample one by one adjusting the weights at each iteration. Thus, the new weights w^* at each iteration are computing solving this quadratic program with linear inequality constraints:

$$min(||w^* - w||_2^2) s.t. \sum w_i^* (T_i - F_i^m)^2 < \sum w_i^* (T_i - F_i^s)^2, \forall s \neq m \tag{2}$$

where w_i presents the current weights of the system, F^m are vectors in the same class as the current training sample, and F^s are vectors belonging to other classes. If a vector of a given class was not already introduced, this step is skipped. For training the face classifier we used a set of 300 images from the internet.

3.3 Classification

For classification we used a simple nearest neighbour approach using the trained weights. One important note is that face types do not always fall completely into one category, as they can be a blend between two different classes. Therefore, we consider the closest two types, and a blending score is computed by simply dividing by their sum. If the scores are within a 60%-40% range we label the face type as a blend and during the recommendation type we allow recommendation from both face types.

4 Eye-ware shape extraction

Unlike the case for faces where we could use an existing face-tracker to extract the geometric features, in the case of the eye-ware frames, accurate extraction of the polygonal shape of the frame is necessary and, as pointed out in the related-work section, no current method is general and robust enough. To extract the polygonal shape of the eye-ware reliably we make three important observations: the eye-ware shape is a closed polygon, it is nearly always convex and symmetric for the right and the left eye [7]. Due to the strong symmetry prior we need to extract only one side.

First the contours are extracted from the image at different threshold levels using a Canny edge detector [4] (Figure 5b). For clarity only one level is shown in the figure. Next, the face boundary line, symmetry line (Figure 5c - dark blue line) and the eye region (Figure 5c - green rectangles) are computed using a state of the art face tracker. Only the contours that intersect the eye region are kept (Figure 5c - purple lines). The contours are further refined by selecting only those that have a symmetric pair (Figure 5d - orange lines). Symmetry is achieved if more than 50% of the points have a corresponding symmetric pair on the paired curve (within a 5×5

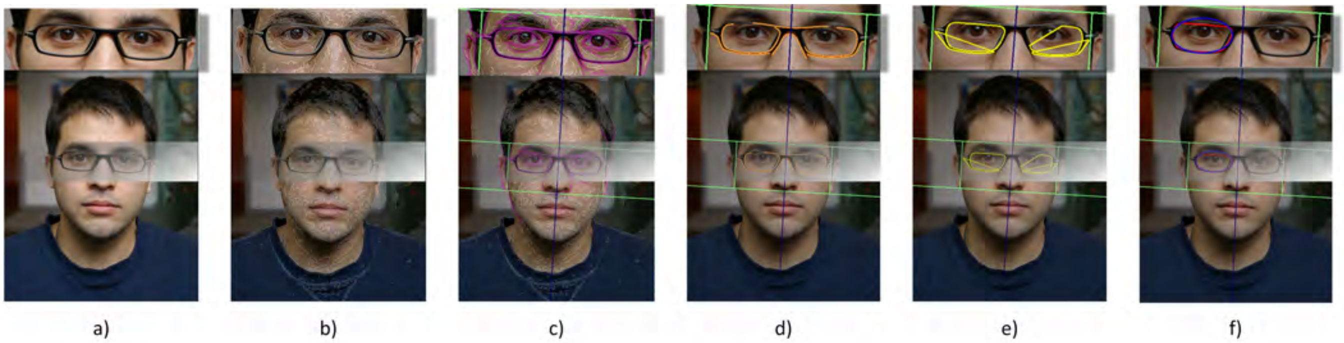

Figure 5: Eye-ware shape extraction. a) Original image. b) Edge detected using a Canny edge detector. c) Large edges that intersect the eye region. d) Contours that have a symmetric pair. e) Convex hull of these contours. f) Selection of the optimal contour (red). A best fitting ellipse to this contour is rendered in blue.

pixel region). Convex hulls are computed on the remaining polygons (Figure 5e - yellow polygons) and are candidates for the final polygon that approximates the eye-ware shape. In order to avoid degenerate cases such as the triangular polygons in Figure 5e. a constraint is placed on both the size of the polygon (more than 20% of the eye region) as well as its roundness (the area of the polygon should be at least 80% of the area of the best fitting ellipse). If more than one candidate exists, the largest one is selected. If no candidates exist, the edge detection threshold is decreased and the process is repeated.

5 Eye-ware shape classification

Once the polygonal shape of the eye-ware frame is extracted, our eye-ware classification follows closely the method of the face classification. For eye-ware, the following classification is widely used in the fashion industry: rectangular, square, round, aviator, wayfarer, oval. These classes are illustrated in Figure 3.

Key geometric features tailored to differentiate between the various eye-ware shapes are extracted from the polygonal shape. Next, these features are converted into a vector that can be used in the CBR framework. This step follows closely the methodology presented in Section 3.2. Finally the classification is based on a simple nearest neighbour search.

5.1 Geometric feature extraction

If we roughly approximate the shape of the eye-ware by an ellipse, the key distinctive features stem from the orientation of the main axis, the size of the ellipse and the degree of symmetry with respect to the horizontal axis (i.e. the eye line). For the aviator and wayfarer eye-ware frames the orientation of the ellipse is somewhat diagonal (Figure 6). For the round and elliptical eye-ware frames the fit of the points to the ellipse is a strong indicator.

Therefore the geometric feature extracted is the best fitting ellipse (Figure 6 - blue polygon) of the eye-ware polygon (Figure 6 - red polygon, the principle axis that shows the degree of rotation are shown in green). Using the best fitting ellipse, a five dimensional feature vector F is computed consisting of the following: the ratio between the height and the width of the best fitting ellipse (F1), tilt angle of the best fitting ellipse (F2), area of the shape polygon as a fraction of the eye region (F3), average distance from the shape polygon to the ellipse (F4) and the distance between the centre of mass of the polygon and the ellipse centre (F5). These features correspond intuitively to discriminating features of the six classes of eye-ware. (F1) and (F5) can discriminate between round against elliptical and square against rectangular. (F2) can discriminate between aviator and wayfarer against the rest. (F3) can discriminate between rectangular and elliptical against the rest (this is because the elliptical and rectangular tend to be smaller than the rest). (F4) can discriminate against round and elliptical against the rest.

Figure 6: Geometric features that we used for the eye-ware shape classification. Left: Sample image with the eye-ware shape polygon extracted (red). Right: Eye-ware shape polygon (red), best fitting ellipse blue), ellipse's axis of rotation (green).

6 Data-driven Eye-ware Recommendation System

In our eye-ware recommendation study we used a database of 240 images collected from the internet and created a survey that asked a set of 42 persons to rate the compatibility between the face shape and the eye-ware shape in each image on a scale from -3 to 3, disallowing a score of 0. Each person voted on about 200 pictures picked at random, grouped in sets of five for each screen. Some images were repeated to check consistency of the choices. The persons received no a priori training and voting on such a large number of images took on average about $20-40$ minutes per person. We had 88 persons participating, collecting overall more than $20,000$ records.

Given this data, the recommendation stage, illustrated in Figure 1, is executed as follows: given an input image of a person not wearing eye-ware, the face type is first classified in one or two of our six classes depending on the face shape. Based on the data collected, we select the best eye-ware type for the face shape and we recommend two frames from an existing database of eye-ware frames.

7 Discussion

We split the discussion section according to the contributions.

7.1 Face shape classification

We trained our classifier on 300 images and evaluated the accuracy of our classification on a different set of 100 images. The accuracy

Figure 7: Examples of mixed shape classes. Left: Half oval and half rectangular face shape. Right: Half oval and half rectangular eye-ware.

of our face classification is about 80%. We performed a t-test to compute the confidence interval: $[74.5\% - 85.5\%]$ with $(p < .05)$ and $[72.7\% - 87.3\%]$ with $(p < .01)$.

Out of the 20% misclassified faces, around half are misclassifies due to partial failure of the face tracker. The classification on a dataset that has correct face extraction is about 90%.

As mentioned before, this classification system is not mutually exclusive, it is possible to have faces and eye-ware frames belonging to more than one class as illustrated in Figure 7. Our system supports these cases.

7.2 Eye-ware shape extraction

Our shape extraction method reconstructs a good approximation of the eye-ware frame shape in 99% of the cases. The validation whether the approximation is correct or not is done by visual inspection. We performed a t-test to compute the confidence interval: $[97.7\% - 100\%]$ with $(p < .05)$ and $[97.2\% - 100\%]$ with $(p < .01)$. Our performance is better than the algorithms presented in the related work section, although it was not done on the same database as there are no publicly available database for this problem. We will make our data publicly available. The closest is the work of Borza et al. [2], but their method is very limited as it requires that the shape of the frame to already be available in a database of reference eye-ware frames. In our method, extracting the shape has no a priori constraint on the eye-ware frame.

Our algorithm has three parameters, as noted in Section 4, that are fixed throughout all our experiments. Out of these parameters the symmetry tolerance needs to be further discussed. Increasing this region will make the algorithm more robust to rotated poses in the images, but it decreases the power of the symmetry prior. Experimentally we chose this region to be 5×5 pixels that worked well in our experiments. However, depending on the input image resolution this might need changing. A better solution might be to normalise the region by the size of the entire image.

7.3 Eye-ware shape classification

We used 36 images of persons wearing frames to train the classifier and 240 to test the accuracy. Our eye-ware frame classification method achieved an accuracy of 90%. We performed a t-test to compute the confidence interval for the eye-ware classification: $[86.8\% - 94.2\%]$ with $(p < .05)$ and $[84.6\% - 95.4\%]$ with $(p < .01)$.

Figure 8: Example of one entry in the validation query: four eye-ware frames were presented for the same person, two recommended by us and two others. Top row shows the system's top two recommendations.

7.4 Recommendation system

To validate our recommendation system, we did a follow-up study. We collected a set of pictures of persons that do not wear eye-ware and using one of the virtual try-on online systems [4] we virtually added four different types of eye-ware frames: 2 were the ones that the recommendation system selected and two selected at random. We selected to show a total of four eye-ware frames for each person because brick and mortar eye-ware retailers mentioned that three to five pairs is what they typically show customers in one batch. We decided to show two recommended pairs instead of just one because, in certain cases the difference in preference between the first and second recommended eye-ware frames can be small. The 42 participants were asked to select their favourite eye-wareframe. Figure 8 shows such an example extracted from our survey. 82% of the times the selection made coincided with one of our recommendations. We performed a t-test to compute the confidence interval: $[77.3\% - 87.7\%]$ with $(p < .05)$ and $[76.2\% - 88.2\%]$ with $(p < .01)$.

8 Conclusion

In this work we introduced a data-driven eye-ware recommendation system based on geometric feature extraction and analysis. At the core of our system are a novel shape extraction method for eye-ware shape and classification methods for both face and eye-ware shapes. The eye-ware extraction method is efficient, more accurate and more robust than previous methods. It works even on difficult input such as sun-glasses with reflection (Figure 9a), partially occluded faces, thin metallic eye-ware frames (Figure 9b) as well as images showing a person with significant pose rotation (Figure 9c). Our classification method uses geometric features tailored for this particular problem in a case-based

[4] http://www.glassesusa.com

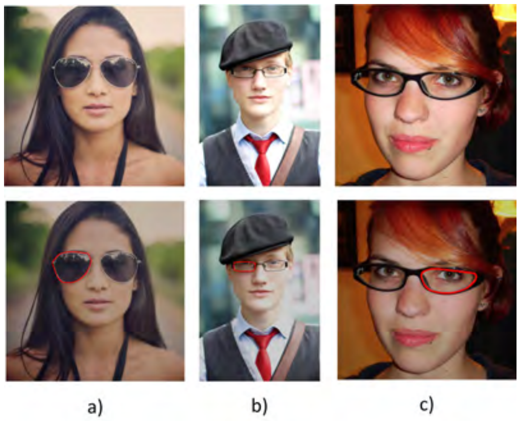

Figure 9: Difficult cases that are solved correctly by our eye-ware shape extractor: a) Sun-glasses with complex reflections. b) Thin metallic eye-ware frames. c) Faces with significant rotations from the frontal position.

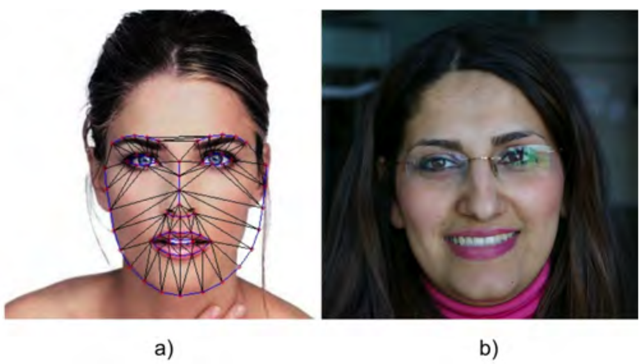

Figure 10: Failure cases: a) Face tracker failure. b) Rimless glasses.

reasoning framework. It provides excellent results despite the small shape variance between classes. Our recommendation system is scalable and efficient, and our validation survey suggests that our system's recommendations are appealing to the users.

8.1 Limitations

Our recommendation method only considers geometric features, although other criteria such as color, style and age criteria can be trivially added. Although our classification mechanism is robust, it can occasionally fail either due to failure of the face tracker (Figure 10a) or failure to correctly detect the polygon when the edges are too faint as it is the case particularly with rimless eye-ware frames (Figure 10b). The face tracker might also fail on images of persons with excessive facial hair.

8.2 Future work

Our work can be extended in a number of interesting ways. Introducing more criteria such as hair type and color, and age and geographical location to the database might improve the quality of the recommendation.

Accurately detecting the shape of the eye-ware frame can lead to methods of removing the eye-ware frames from a photograph or video stream. Also using the robust eye-ware frame shape extraction, our method can be extended with a smart browsing feature where the user can explore similar shapes to the ones recommended based on shape similarity.

ACKNOWLEDGEMENTS

This research was funded in part by the Natural Sciences and Engineering research Council (NSERC) of Canada and ENCS faculty, Concordia University, Montreal.

REFERENCES

[1] A. Aamodt and E. Plaza. Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI communications*, 7(1):39–59, 1994.

[2] D. Borza, A. S. Darabant, and R. Danescu. Eyeglasses lens contour extraction from facial images using an efficient shape description. *Sensors*, 13(10):13638–13658, 2013.

[3] A. Bottino and A. Laurentini. The analysis of facial beauty: an emerging area of research in pattern analysis. In *Image Analysis and Recognition*, pages 425–435. Springer, 2010.

[4] J. Canny. A computational approach to edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (6):679–698, 1986.

[5] B. Fink and I. Penton-Voak. Evolutionary psychology of facial attractiveness. *Current Directions in Psychological Science*, 11(5):154–158, 2002.

[6] X. Jia and J. Guo. Eyeglasses removal from facial image based on phase congruency. In *Image and Signal Processing (CISP), 2010 3rd International Congress on*, volume 4, pages 1859–1862. IEEE, 2010.

[7] Z. Jing and R. Mariani. Glasses detection and extraction by deformable contour. In *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, volume 2, pages 933–936. IEEE, 2000.

[8] M. Lai, I. Oruç, and J. J. Barton. The role of skin texture and facial shape in representations of age and identity. *Cortex*, 49(1):252–265, 2013.

[9] T. Leyvand, D. Cohen-Or, G. Dror, and D. Lischinski. Data-driven enhancement of facial attractiveness. *ACM Transactions on Graphics (Proceedings of ACM SIGGRAPH 2008)*, 27(3), Aug. 2008.

[10] I. D. Papel. Computer imaging for facial plastic surgery. *Facial plastic and reconstructive surgery (2nd edition), Thieme Medical Publishers, Inc, New York*, pages 110–115, 2002.

[11] J.-S. Park, Y. H. Oh, S. C. Ahn, and S.-W. Lee. Glasses removal from facial image using recursive error compensation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(5):805–811, 2005.

[12] C. K. Riesbeck and R. C. Schank. *Inside case-based reasoning*. Psychology Press, 2013.

[13] J. M. Saragih, S. Lucey, and J. F. Cohn. Face alignment through subspace constrained mean-shifts. In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 1034–1041. IEEE, 2009.

[14] S. Schönborn, A. Forster, B. Egger, and T. Vetter. A monte carlo strategy to integrate detection and model-based face analysis. In *Pattern Recognition*, pages 101–110. Springer, 2013.

[15] J. Shotton, A. Blake, and R. Cipolla. Multiscale categorical object recognition using contour fragments. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(7):1270–1281, 2008.

[16] K. B. Sun and B. J. Super. Classification of contour shapes using class segment sets. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 2, pages 727–733. IEEE, 2005.

[17] H. Wu, G. Yoshikawa, T. Shioyama, S. Lao, and M. Kawade. Glasses frame detection with 3d hough transform. In *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, volume 2, pages 346–349. IEEE, 2002.

[18] Y. Xiao and H. Yan. Extraction of glasses in human face images. In *Biometric Authentication*, pages 214–220. Springer, 2004.