MediaTek Video Face Beautify

November 2014

Table of Contents

1 Introduction

Improved camera-equipped mobile device quality allows device users to populate their diaries and social media postings with their own photographs. And some users apply software-aided manipulation to these photographs to enhance the beauty of these images, particularly facial images.

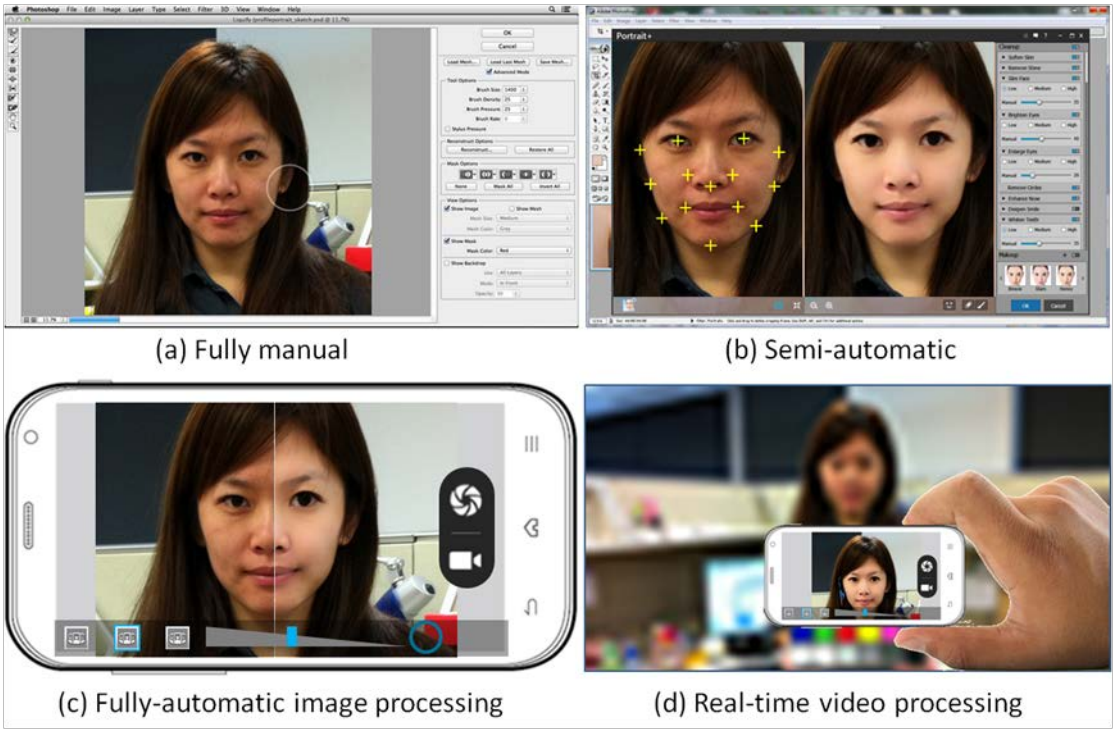


Figure 1. Face Beautify Progression

"Face Retouching" or "Face Beautify" technology refers to multi-step image processing to adjust the color, texture or shape of human faces. The goal can vary, such as removing wrinkles and blemishes, lightening skin, and enlarging eyes.

In the past people manually modified photos, sometimes using complex image processing software (Figure 1a), which sometimes requires considerable software skills. With improved face beautification technology device users can beautify those faces by selecting several facial features and one target template (Figure 1.b). Today advanced facial feature algorithms enable fully-automatic face beautification. In this fully-automatic process the user has to capture a facial image and the face will be processed automatically within seconds (Figure 1.c).

Undoubtedly, the best user experience would be "What you see is what you get," with zero time processing delay (Figure 1.d). However, due to algorithm complexity and the limitation of computation resources, applying in mobile devices face beautification effects on the real-time preview screen remains difficult.

2 The MediaTek Solution

Mediatek has developed the **Video Face Beautify** application which seamlessly combines advanced computer vision algorithms and all resources, including CPU, GPU and HW, in the mobile device to achieve real-time face beautification. In this application, four core functions are included to perform facial beautification: wrinkle and blemish removal, skin tone adjustment, eye enlarging, and face sliming. With the benefits of this technique, the users can adjust the parameters and see real-time results for better face quality. Additionally user can capture photos or record videos for later social sharing. To our knowledge, this is the first solution in the world which meets all the targets of face beautify (including face reshaping) in all camera modes (preview, capture and video record).

We add the term "Video" to the commonly used industry term "Face Beautify" (vFB) to differentiate our solution from those of competitors. In the following sections, we will first briefly introduce the vFB system flow and function units. The main techniques used in this system include face detection, skin region detection, color adjustment, and shape analysis for non-rigid face warping.

3 Overview of Video Face Beautify

This section introduces a global view of the vFB algorithm. The system flow is defined based on the combination of advanced computer vision technologies and each system block is optimized after trying different methods. Since vFB is developed for real-time applications, we integrate software and hardware for flexibility and performance. The vFB system can be simply grouped into three layers (Figure 2):

1. Computer Vision Technologies
    - To achieve the vFB implementation target, four important vision algorithms will be used in this system: face detection, skin detection, color analysis, and shape analysis.
2. Basic Algorithms and Model
    - Some general algorithms are commonly appeared in different high-level computer vision research topics, such as off-line trained model, optimization and classification tools. These algorithms are modified and improved for MTK's vFB application.
3. Optimized Task Scheduling
    - In this project, the resources of CPU, GPU and HW are integrated for best performance consideration. Therefore, advanced optimization algorithm is important for optimal task scheduling.

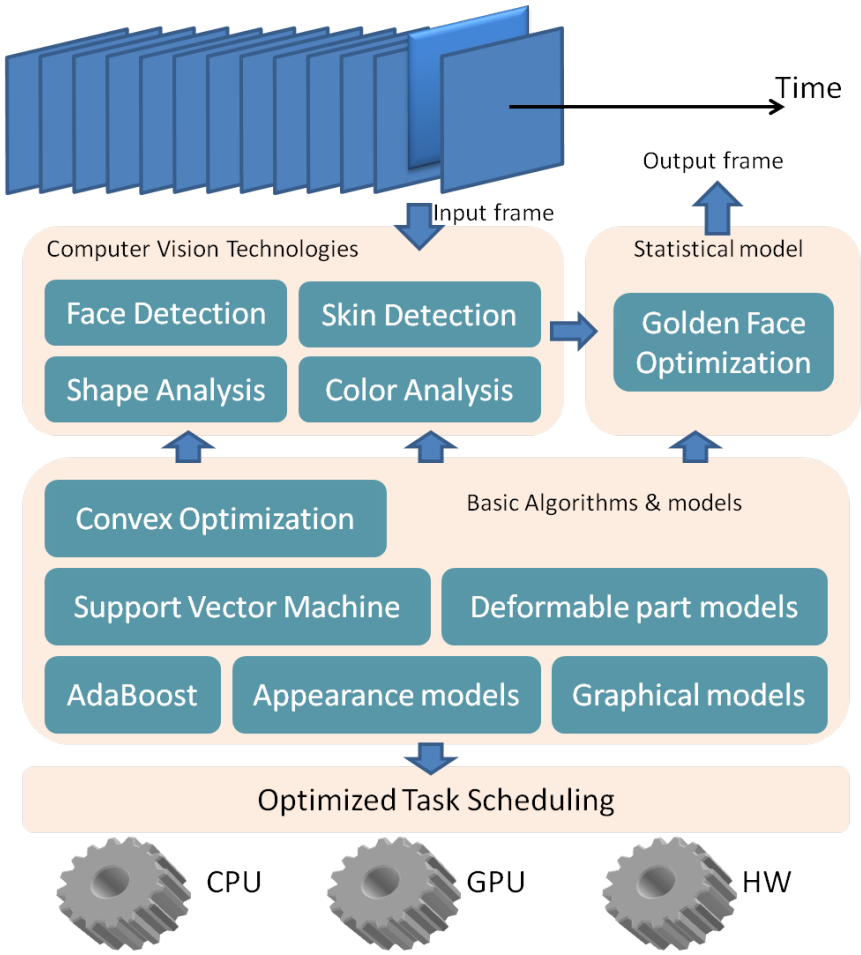


Figure 2. vFB System Overview

With these three layers, the benefits of our solution include computational efficiency and significant flexibility.

The core comprises the following technologies: **Face Detection, Skin Detection, Shape Analysis and Color Analysis.**

- The **Face Detection** unit extracts the face location and orientation. Several basic vision algorithms, such as SVM and AdaBoost, are used in this unit.
- The **Skin Detection** unit is designed to remove wrinkles and blemishes, and to adjust the facial skin color. The **Skin Detection** unit detects the skin area by using off-line trained skin-color models and dynamic skin classification algorithms.
- The **Shape Analysis** adjusts the facial shape, like enlarging eyes and sliming face. The input facial shape is analyzed and fit to our target golden face in the **Shape Analysis** unit.

- The **Color Analysis** unit utilizes skin area information from the Skin Detection unit to analyze skin color. The skin color can be adjusted properly according to original facial skin color for more natural results.

4 Face Detection

Face Detection, widely used in many devices, aims to detect face location and orientation under wide variations in pose, scene lighting, and facial appearance. In the MediaTek solution we carefully design the face detection unit in the detector architecture and in the underlying training methodology employed to obtain classifiers used by the detector. Our streamlined detection framework can be used as a basis for future improvements.

Similar to traditional face detection algorithms, this unit uses scanning window detection, boosted cascades and integral images for fast image feature computation. However, the computational efficiency in our detection architecture differs from traditionally adopted methods. First, we use modified haar-like features, and tree-structure searching strategy to limit lighting variation in environment. Secondly, this unit employs a temporal-shift search strategy to further speed up detection.

To improve the detection accuracy and coverage, we employ a radically different detector training methodology. We built a large-scale training set by sampling images captured using smartphone cameras. Furthermore we adopted a confidence-weighted boosting procedure to build tree-structure classifiers. In addition, this unit uses a modified version of the multiple instance pruning framework in order to compute the optimal set of decision thresholds. And we can obtain the face location and orientation accurately in real-time from the Face Detection unit.

5 Skin Detection

The skin region can be determined by computing skin probability from different color channels.

(a) Original

(b) Skin probability in each color channel

(c) Skin area

Figure 3. Skin Detection

In this application, the skin region is estimated in a probability computation based considering the color distribution, statistical information and feature geometry, the optimal skin region can be finally determined.

To find the skin region, we compute the skin color probability in each color channel (figure 3.b). With a training process and online color analysis, we obtain an optimal weighting for each channel. However, color may not always work to determine skin region. Therefore, an additional statistical facial mask is also used during computing the probability of skin pixels.

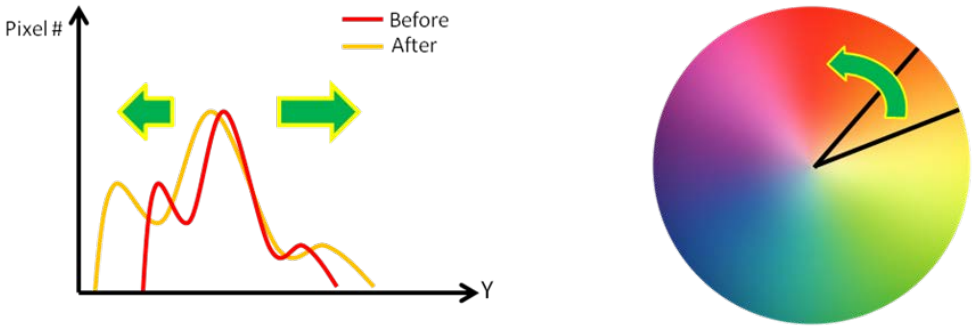
5.1 Wrinkle and Blemish Removal

Based on the detected skin region (Figure 3.c), we first remove blemishes on the face, such as a mole. After blemish removal, the wrinkle on face can also be removed by applying edge-preserving smoothing on each pixel. This step allows us to remove facial wrinkles and to make the skin appear smoother. Note that we ***only*** smooth the facial skin region; other regions, including the background area, are ***not*** modified.

Directly smoothing the image globally without skin detection, as performed in some 3rd party solutions, will blur background details. This leads to poor video quality.

5.2 Skin Tone Adjustment

For skin color adjustment, the brightness is analyzed automatically and the color of the skin will be adjusted according to the facial skin tone by multiplying a scalar to enhance the brightness and contrast of face. Since the skin region is accurately detected in our solution, the skin smoothness and color is adjusted dynamically in a natural way (Figure 4 below). Similar to the description in previous sub-section, many other competitor solutions adjust the brightness or color in the whole image, which may lead to over saturation or false background color. The background or non-skin region is not affected in Mediatek's solution.

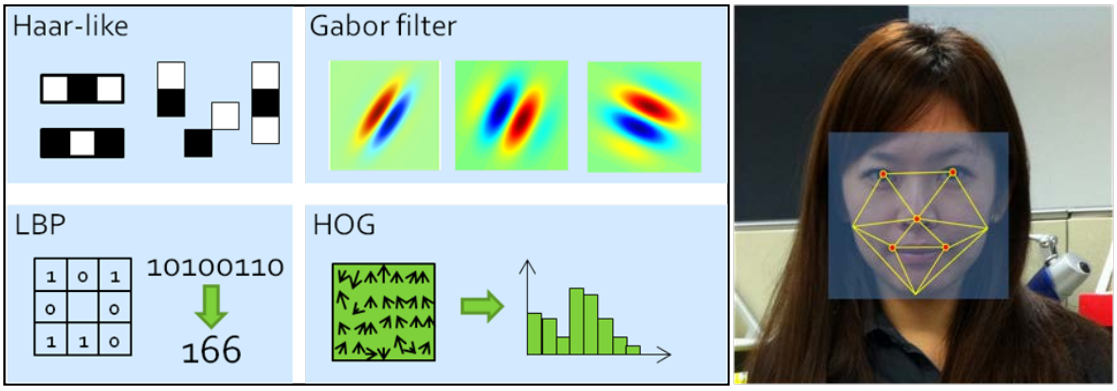


(a) Enhance Contrast & Brightness (b) Adjust Skin Tone

Figure 4. Skin Color Adjustment in the MediaTek Solution

6 Shape Analysis

Based on the Face Detection algorithm, we also developed a fast and more accurate shape analysis algorithm.



(a) Appearance Model (b) Shape Geometry Model

Figure 5. Appearance and Shape Geometry Models for Shape Analysis

Based on the Face Detection algorithm, we also developed a faster and more accurate shape analysis algorithm. Several low-level features can be used as **Appearance Model**, like Haar-like feature, Gabor filter, Local Binary Patterns and Histogram of Oriented Gradient. (Figure 5.a)

Since the classifier is applied to the output of the Facial Shape Analyzer in a sliding window fashion, we need to reduce the search range for score computing. Therefore, all the facial features are connected and the covariance of each component is restricted by a graph-structure. This graph-structure defines the **Shape Geometry Model** and facilitates our finding the optimal combination between features. (Figure 5.b)

6.1 Golden Ratio Facial Proportions

To better understand common notions of facial beauty, we analyzed more than 12k facial images of different races. In addition, we aligned all the face images and computed an average face as shown above in Figure 1. The average face provides information about the ratio between facial features, which is very similar to the so-called golden ratio proportions.

Based on our user study and facial proportion analysis, we can target the input face to the ideal golden one. In order to better adjust the shape of input face, we need to accurately detect the facial feature parts and automatically analyze the skin condition and the proportions between features. (Off course, users can also adjust all the parameters for preference with UI control

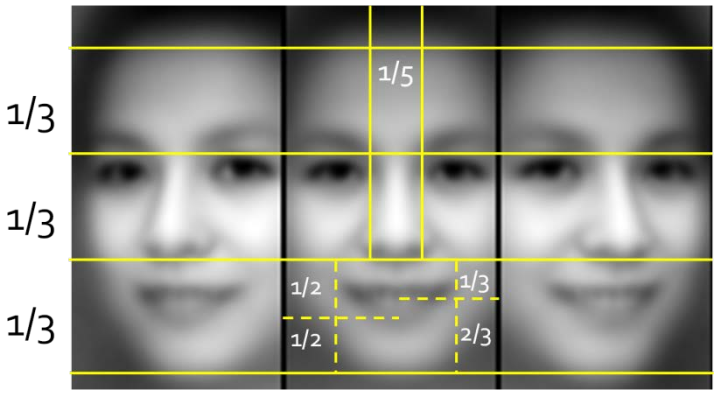
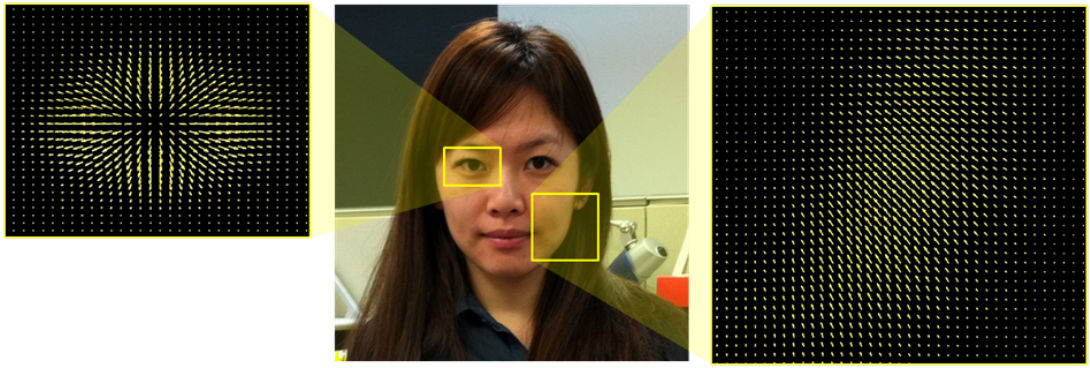


Figure 6. The Average Face and the Golden Ratio

6.2 Facial Shape Optimization

Based on the user studies and detected facial features, we are able to compute the difference between input face and the target golden one. And we locate the center of eyes and the size of face for advanced shape adjustment. Please see Figure 1 for an example, the movement of every pixel in each feature component will be estimated based on the feature geometry analysis, including moving direction and intensity. If the warp intensity is too weak, the face shape will not be adjusted. On the other hand, if the warp intensity is too strong, the personal style of the face will miss. Therefore, face optimization process aims to find the optimal value for best adjusting the face shape toward the golden one and keeping personal style simultaneously.



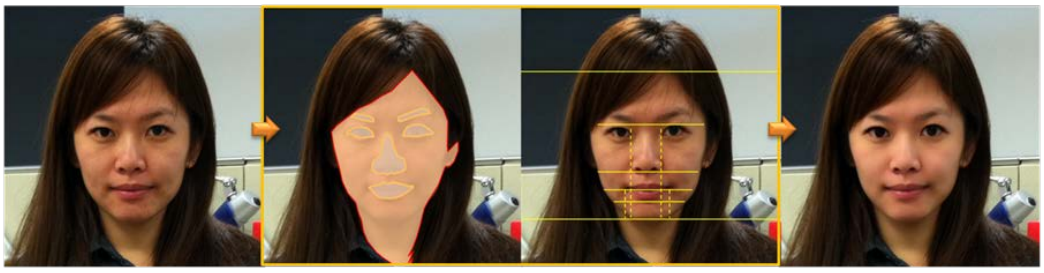
(a) Pixel flow around eyes (b) Input face image (c) Pixel flow around face

Figure 7. Pixel Flow for Optimal Face Shape Adjustment

As long as the warp information is estimated, all the pixels are automatically relocated to new positions.

7 Conclusion

Mediatek's Video Face Beautify solution combines cutting-edge technologies from computer vision and making real-time intelligent face beautification possible. With software optimization, we are able to locate faces, analyze skin color and facial shapes within 30 milliseconds.



(a) Original Image (b) Skin & Color (c) Shape Analysis (d) Face Optimization

Figure 8. The Face Beautification Process

With optimized task scheduler, we can seamlessly integrate the resources of CPU, GPU and Hardware for best quality output (1080p). On the other hand, we provide a user friendly UI for the users to adjust the parameters of all the functions and see the beautified results on the screen simultaneously.

To the best of our knowledge, this is the first solution in the world which meets all the targets of face beautify for high resolution output (Full HD) in all camera modes (preview, capture and video record).

In the future, we will keep working hard to powers the mobile devices and anyone can achieve something amazing with the right help, just like Video Face Beautify.