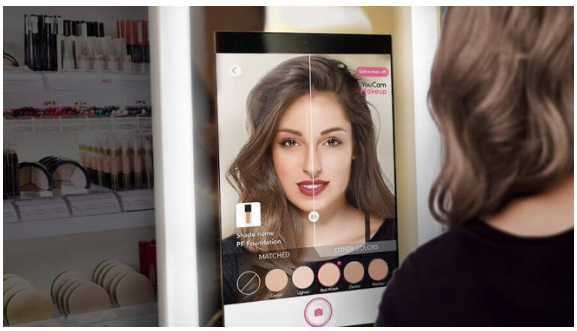
How AI and AR can help beauty industry

Johnny Tseng, Ph.D.
CTO, PerfectCorp
johnny@perfectcorp.com

Honder Tzou
Chief Architect, PerfectCorp
honder@perfectcorp.com



Cosmetics and humans have a very long history. Egyptians, 7,000 years ago, saw beauty as a sign of holiness—and used copper carbonate to create green eyeshade. To the Greek philosopher Plato, beauty was indistinguishable from philosophy. Ancient China pioneered painting fingernails. Tribes in Africa—as elsewhere—have used face paints and other elements of cosmetics to denote affiliation, roles, and status. Across the years, poets, painters, and philosophers have weighed in on the role of beauty in cultures around the world. And virtually everywhere, cosmetics in one form or another—for men and women—have played a role. A few years ago, ELLE magazine, in a photo essay “21 Portraits of Beauty Around the World” [1] provided a nice definition by the Romanian photographer Mihaela Noroc, who said: “Beauty is everywhere. ... We are all beautiful because we are different." It’s only natural that the world of high technology would soon find a role in beauty — especially with tech’s innovations in artificial intelligence (AI) and augmented reality (AR). These technologies are helping to democratize beauty and help each person find the best product for their unique features.

One could trace the birth of Beauty Tech back at least to early paint and photo shopping applications on desktop computers, though it really took off with the more recent advent of AI and AR technology. “Although beauty tech dates back to the days of boxed computer programs (Cosmopolitan magazine put its name on “virtual makeover software” all the way back in 1998), it’s only recently that big brands have started to invest heavily in the space. The surge in interest has been driven in large part by advances in smartphone technology and the development of advanced beauty tech solutions such as Perfect Corp.’s AR virtual try-on solution for makeup and AI skin diagnostic technologies for skincare brands. Consumers now have their smartphone and through the power of beauty tech, can snap a photo to try on makeup shades and receive a personalized beauty recommendations in seconds.

Virtual Try-On

One of the perhaps underappreciated wonders of AR is that it allows users to do what would simply be impossible to do otherwise. Estée Lauder’s first implementation of virtual try-on solutions provided by PerfectCorp offered customers the magic of being able to try “30 Lipstick Shades in 30 Seconds.” Presently, lipstick aficionados can virtually try on every shade in Estée Lauder’s Pure Color Envy lipstick range with the brand’s Virtual Artist Tool. In real life, trying 30 shades in even 30 minutes would be unlikely, as the lips would become too tender from the cycle of applying and eventually affect the applied result. The same is true for just about any cosmetic—and even more so with something as time-consuming and as difficult to reverse as dying your hair.

To accomplish virtual try-on task in a device is not an easy task. It contains facial tracking, understanding the lighting environment and makeover rendering in a video frame cycle which has to be done in less than 50ms. The target device can be a mobile phone, a tablet or a browser which computation power is very limited. Moreover, the virtual try-on solution has to simulate not only single effect (e.g. lipstick) but all the combinations (e.g. a combination of lipstick, eyeshadow and foundation) in a cycle. Due to the limited environment, all the computation tasks have to be fully optimized to achieve the goal.

Facial Tracking

A fully immersive AR experience depends upon faultless facial tracking. Users won’t sit still when experiencing the virtual try-on. They want to move their head, similar as they would in front of a mirror. The lack of faultless facial tracking can result in lipstick lags behind when moving from side to side, or that the AR applied hair color gets stuck on the wall when the user moves to see the AR effect from different angles. An enormous amount of technology goes into real-time facial tracking. Without this technology, makeup can take on a clownish appearance, and the session, while perhaps a bit amusing, won’t bring satisfaction nor bring users back. Perfect Corp.’s AgileFace® tracking technology accurately detects all facial features with the utmost precision. This technology properly identifies all variations of facial features, resulting in the most inclusive virtual try-on experience.

AgileFace is a fully optimized facial tracking engine which output 84 points of facial landmarks at 100 fps on iPhone 7. The inter-ocular NME (Normalized Mean Error) on our test set is less than 3.5%. Surprisingly, the model size is only 942KB. Figure 1 shows the training pipeline of Perfect Corp’s AgileFace. At beginning, several tens of thousands of facial images with landmarks are carefully prepared.3D facial model is selected from the 3D facial expression database for each face within the image. The model is then rotated

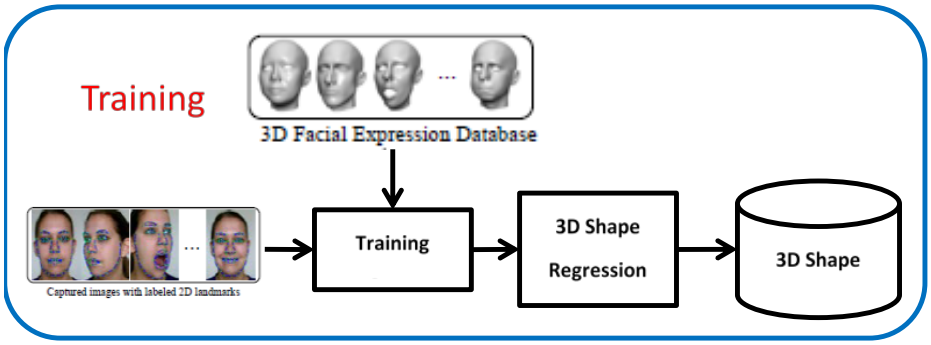


Figure 1: AgileFace training pipeline

and transformed to fit into the 2D landmarks. The fitting error is minimized through a regression process. Afterwards, all the data are fed into network for regression training.

Figure 2 shows the inference process in the run time. Image with a face is fed into the inference engine and the output is a 3D facial model with 3D facial landmarks, expression and transform parameters. The data together with the original image are fed into the refinement network to refine the landmarks for better estimation. The refinement network references the data of the previous frame and the current data for better prediction.

Lighting Environment Estimation:

Color derives from the spectrum of light (distribution of light power versus wavelength) interacting in the eye with the spectral sensitivities of the light receptors. Color is also associated with objects or materials based on their physical properties such as light absorption, reflection and lighting emitted from the light sources. For virtual try-on, from the user perspective, the color is received from the display instead of the real materials. The color in the display is not related to the absorption, reflection and the environment light sources. Instead, the color is emitted directly from the LED (light-emitting diode) to the eyes. The virtual try-on process is constructed by 2 phases. The first phase is to receive the color from the materials interacted with the surrounding lights and then received by the camera. The second phase is to blend the virtual materials (e.g. eyeshadow) interacted with the same lighting conditions and then present the mixed image on the display. The material parameters (or shades) are pre-defined and the only unknown parameter is the lighting environment. Through simple deep learning model with various tagged data images, the lighting parameters can be easily predicted.

Rendering

Cosmetics are more complex than what you think. There are lots of cosmetic types such as lipstick, lip gloss, lip liner, eyeshadow, blush, mascara, eye liner, foundation, bronzer, highlighter, concealer, etc. Cosmetics are constituted from a mixture of chemical compounds derived from either natural sources, or synthetically created ones. The shade is kind of complex and is not just a single color. For example, besides the color, a lipstick may have one or more of different material types such as matte, satin, sheer, glossy, shimmers, metallic and holographic. Figure 3 shows the virtual try-on results of different lipsticks. Besides precise colors, virtual try-on mixes the shades together with the lip texture. You may find that, some lipsticks may conceal the lip-lines, some enhance the lip-lines and some conceal the thin lip-lines but enhance the think ones. If lipstick is glossy, or metallic, shimmers, or

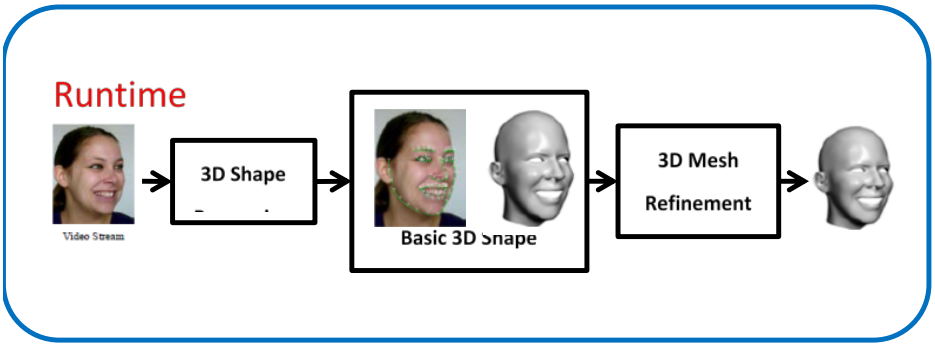


Figure 2: AgileFace inference process in run time

has other unique features, the light reflection results are quite different. Precise blending is required to enable the user to see exactly how it will look on their own lips. Failure at this results in unhappy customers and can have devastating effects on social media. By migrating this hyper-realistic virtual try-on experience on website, customers can instantly try on dozens of popular makeup products, including lipstick, eye shadow, mascara, highlighter, and more. The digital effects of a virtual makeup try-on so closely mimic those of a physical try-on that consumers can be confident they are making the right purchase decision.

Figure 3: Virtual try-on results by different lipsticks

Skin Tone

Among cosmetics, foundation is the basis of the whole makeup on a face. Foundation is a liquid, cream, or powder makeup applied to the face to create an even, uniform color to the complexion, cover flaws and, sometimes, to change the natural skin tone. Besides foundation, contour and highlighter have the similar ingredient but different purposes. While foundation creates the base for the makeup, contouring makes the face look chiseled and sharp, and highlighter highlights the highest points of your face. Foundation covers all the imperfections like acne, pigmentation, redness, dark spots etc. giving the face an even tone. Contour gives you amazing cheekbones, jawline and hairline, pointed nose by darkening the areas of the face where a shadow would naturally fall. Highlighter, on the other hand, enhances the effect of contouring by highlighting the areas on which light would hit, like the highest points of cheekbones and brow bones, forehead, and nose bridge. Figure 4 shows the foundation, contour and highlighter result through PerfectCorp's virtual try-on technology.

Figure 4: Foundation, contour and highlighter simulation result left: before, right: after

Usually, the foundation shade has to match the user's skin tone. Foundation is often puzzling for users. Just which foundation is right? Finding the perfect foundation shade can be a challenging process for many consumers, but AI beauty tech can simplify the process. For example, Estée Lauder's latest in-store beauty app combines precision foundation shade detection with a virtual try-on-color match experience. With iMatch™ Virtual Shade Expert, Perfect Corp.'s AI deep learning for real-time skin-tone detection uses its knowledge base of 89,969 skin tones to help customers find their preferred foundation shade in Estée Lauder's number one foundation, Double Wear Stay-in-Place Makeup, and see it virtually applied in real-time. "We found a customer's perfect foundation shade match is not always her preferred shade," says Gen Obata, former Vice President Global Retail Experience at Estée Lauder. "So iMatch™ Virtual Shade Expert answers this need. It finds her perfect shade and then provides options that are lighter, darker, cooler, or warmer so she can choose the shade she thinks suits her best." This kind of personalization creates a strong customer bond.

PerfectCorp's skin-tone detection engine is based on a VGG [2] like backbone with attention model. The training data consists of 20K tagged images and 20K of augmented images. The augmentation is based on 3D based simulation with different lighting conditions. The model size is about 1.1MB and the inference speed is less than 3 seconds on iPhone 7. According to the testing result and based on Delta E function, the color difference between PerfectCorp's skin-tone estimation and the value of a spectrometer is about 1.7, which is less than the JND (Just noticeable distance, 2.3) of Delta-E [3]. This also implies that this software solution is almost the same as an expensive spectrometer.

Hair-Dye

The value of a virtual try-on experience—for the customer as well as the seller—is to provide an authentic presentation of how a product will actually look on the customer. This can be especially challenging with hair color, as the color applied will differ according to the customer's natural hair color. "Lashes are extremely hard to capture in AR," says American International Industries' Marwan Zreik. "They can look pasted on. And the same is true for hair, which is also a major technical challenge to precisely render. We were all blown away by the accuracy of Perfect Corp's AI and AR. The color went onto the hair in a completely natural way. It actually works with the tone of your natural hair color, which means that each of our Punky colors looks different on every person depending on their base hair color." Figure 5 shows the hair-dye effects based on PerfectCorp's hair-dye AR engine. Similar to cosmetic AR, besides the major color, hair-dye supports lots of effects such as ombre, metallic, pigment, shining, half-dye, multi-color which support most of the products in the market.

Figure 5: Hair-dye AR: Ombre, Metallic and Pigment effects.

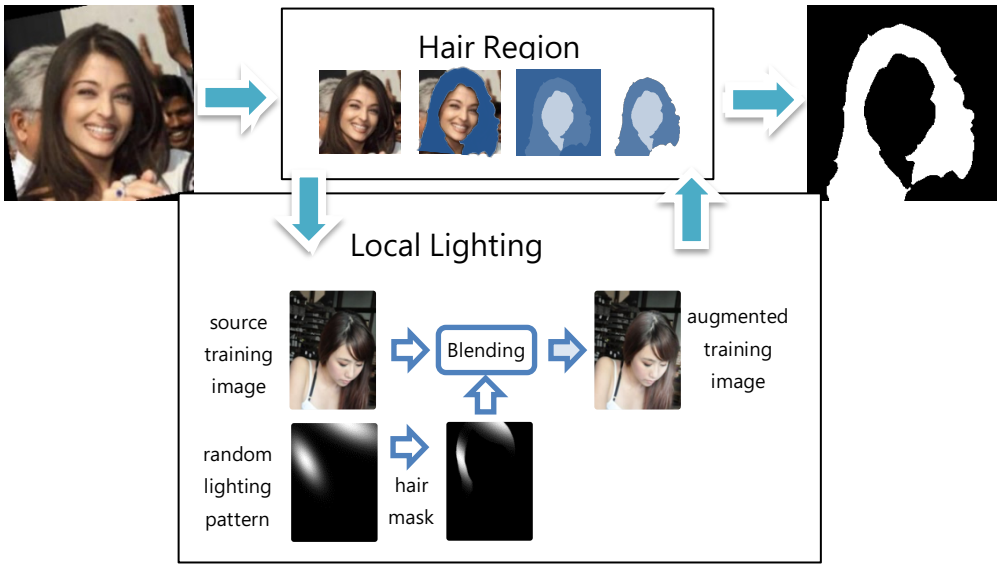


Figure 1: hair-die training process

Figure 6 shows the training process of hair region segmentation. This is the major part in hair-dye AR module. The training data is about 10K tagged images. Lots of (about 10x) light simulation augmented images are used additional to the tagged images. The net is a combination of U-Net [4] and DenseNet [5]. The inference time is less than 15ms on iPhone 7 and the IOU for our test set is about 84%. Our test is a collection of hard samples which contains different lighting conditions, similar color of background and cloth, motion blur and different head angles.

Skin Care

High-definition cameras combined with AI-powered processing allow consumers to use their own smartphone as a skin diagnostic tool accessible anytime, anywhere. A typical use case is offering an online 'skin doctor' experience—accessible 24/7— that can assess one's skin concerns and grade them professionally in terms of severity in real-time. The AR technology can then create multiple layers of overlay images to show the exact location of wrinkles, fine lines, spots, uneven texture, and acne. For the very first time, consumers can receive a precise objective evaluation of their skin condition without visiting the dermatologist. Brands can integrate these tools across their digital channels to provide customers with a skin assessment and customized product regimen based on those results. With Perfect Corp.'s AI Skin Diagnostic Technology, consumers can obtain a personal skin analysis across 14 skin concerns, including: Spots, Wrinkles, Moisture, Redness, Oiliness, Acne, Texture, Dark Circles, Eye bags, Skin Firmness, Droopy Upper Eyelid, Droopy Lower Eyelid, Radiance, and Visible Pores.

With AI skin analysis, it is extremely important to ensure that the diagnostic results are verified by dermatologists and skincare experts. Thus, Perfect Corp. also collaborated with experts in the field to ensure the AI skincare technology can provide accurate, meaningful, and validated results to skincare shoppers.

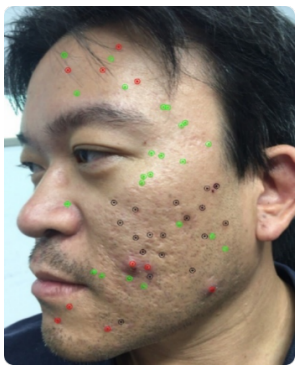


Figure 7: Acne detection, green: comedonal acnes, red: pimples, black: hyperpigmentation acne or acne scars

Figure 7 shows the acne detection result. Acne detection is one of the 14 skin concerns supported by Perfect Corp.'s AI Skin Diagnostic Technology. The marker in green indicates comedonal acne which is normally referred as blackhead or whitehead. The marker in red indicates a pimple which is either a pustule, a papule, a cyst or a nodule. The market in black indicates a hyperpigmentation acne or an acne scar. The net is U-Net like backbone. The most important part is the images and the tags. PerfectCorp work with tens of experts to collect the 10K images and refine the ground truth back and forth. Of course, data augmentation is needed for this case. The inference time is about 3 seconds on iPhone 7.

Recommendation

Today, beauty is not one size fits all, and when consumers shop for new makeup products, they are seeking out personalized advice and recommendations from brands. The power of AR and AI technology for beauty brands lies with its ability to cater to individual consumers and their specific beauty needs. The deep learning algorithms that drive AR and AI beauty tech help brands provide mass scale personalization to consumers. With these technologies, every customer can try-on products virtually and receive a personalized product regimen.

According to a recent study by Forrester, 77% [6] of consumers have chosen, recommended, or paid more for a brand that provides a personalized service or experience. Forbes also recently reported that 71% of consumers feel frustrated when shopping experiences are not personalized. From this data, it is clear that beauty tech tools and personalization strategies are becoming essential to create strong customer experiences and shopping journeys. Figure 8 shows PerfectCorp's Face AI Engine which can detects tens of facial attributes including face shape, eye shape, eyebrow shape, lip shape, nose shape and cheekbone shape. Together with skin tone, eyebrow, eye, hair and lip colors and beauty advisor's recommend-ation rules, the face recommendation engine can provide fully personalized makeup recommendation for users.

Conclusion

As the world emerges from the pandemic, the beauty industry is experiencing a digital transformation. More than ever before, consumers are turning to brands for digital services and experiences

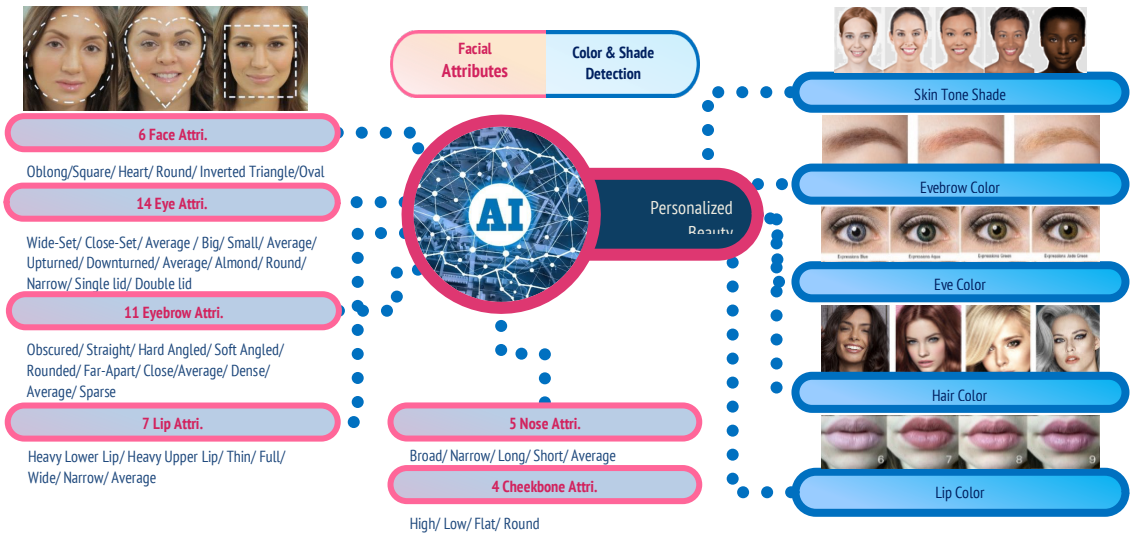


Figure 8: PerfectCorp's Face AI engine which can detect tens of facial attributes and colors.

that offer personalized product advice and inspiration. In order to effectively engage consumers and provide impactful shopping experiences, it has become essential for brands to create digital experiences that entertain shoppers and help each customer find their perfect product match. Beauty Tech, is an emerging new tech trend encompassing AI and AR-powered virtual try-on technologies and diagnostics. Beauty tech has already helped hundreds of beauty brands grow their sales figures and increase customer engagement.

References

[1] E. GRIFFIN, 2015. [Online]. Available: https://www.elle.com/culture/news/a26870/atlas-of-beauty-photos/.

[2] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," in ICLR, 2015.

[3] B. Fraser, C. Murphy and F. Bunting, Real World Color Management: Industrial-Strength Production Techniques 2nd Edition, 2004.

[4] O. Ronneberger, P. Fischer and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2015.

[5] G. Huang, Z. Liu, L. v. d. Maaten and K. Q. Weinberger, " Densely connected convolutional networks," in IEEE Conference on Computer Vision and Pattern Recognition, 2017.

[6] Forrester, "slideshare," 24 2 2017. [Online]. Available: https://www.slideshare.net/TrackIF/forrester-webinar-individualization-versus-personalization. [Accessed 4 1 2022].