DESCRIPTION

*Face Swapping: Automatically Replacing Faces in Photographs* [Bitouk et al. 2008] describes a procedure for automatically replacing the face of a target individual with that of someone else. In other words the eyes, nose, mouth, and etc. of the target’s face will be replaced by that of another individual. These features are modified to match the skin tone and structure of the target individual so that it appears that these are the target’s natural features. The essence of the procedure is to search for candidate faces from a stored library that are most similar in alignment to the target, blend the candidate faces to be visually similar to the target, and finally rank the candidates by similarity.

The primary motivation was to provide an alternative to existing de-identification techniques used for privacy, such as blackout and blurring. The former is an easily noticeable modification of the image, while the latter, although potentially more subtle, also detracts from the image quality by introducing an inconsistency in the sharpness of detail across the image. The example application of face swapping for Google Street View would surely impact the millions that use the service, but the value would arguably be marginal. Street View is used to preview locations, not view to the individuals in that happened to be at the location by chance when the Street View image was taken. The individual’s presence is inconsequential and thus a blurred face in contrast with a swapped face would have little effect on user experience.

The face swapping algorithm has significant impact on fields in which face swapping is performed by automating a process that before required initial user input to select and align images. It also brings more efficiency to the process by working in a simple 2D coordinate system instead of a 3D coordinate system for which factors like angle to the subject must be taken into account.

CLARITY OF EXPOSITION

The paper well outlines the motivation for the development of the procedure, giving an overview of the procedure itself and stating the insufficiencies of existing techniques. The reader is primed by an intuitive example of an application with the widely used service Google Street View, which illustrates the point of security through image modification. Descriptions of an application of a face swap are given at a fairly high level with some technical detail, but nothing too specific to computational photography such that a reader with a scientific background could comprehend the concepts. There are two considerations which could have improved the clarity of this section. It would have been helpful to have the algorithm parameters defined – the pitch, yaw, etc. of an image is not immediately obvious. Secondly, a coordinate system with respect to an image could have benefitted from an example or could have been left out altogether at the level of detail of the exposition without detracting from the reader understanding the process.

QUALITY OF REFERENCES

­­Three references are made to previous work with face replacement. All three efforts work in a 3D coordinate system, in contrast to the method of suggested by the paper, which is in a 2D coordinate system. Additionally, the previous works require manual user input to get the processes underway, while the paper has an algorithm that sets forth automatic replacement. Although contrasting to previous techniques is useful to specify what improvements have been made, the three part listing made is highly redundant and one citation would have sufficed. Furthermore, the authors note for one of the three references, XiD Technologies, that the details of the process used, the extent to which it is automatable, and quality of results are not known. It seems that the effort dedicated to noting this previous work and effort required on the reader’s part to keep track of all works could have been better utilized by excluding this example if there is not more content to say about it than knowledge of the work’s existence. It would have been worth noting if any work had been done in image component replacement in a 2D coordinate space, even that outside of the domain of face swapping. Work in this space would likely be similar to the techniques described in the paper or, if they were not, a contrast to the approach needed specifically for face swapping would have been stimulating. All other references seemed relevant in either citing previous work or specifying techniques that could be used to make the face swapping algorithm more powerful.

REPRODUCABILITY

The steps taken to transfer a face from one individual to another are thoroughly detailed in the paper. In several instances the most intuitive approach for solving a step of the problem is presented, followed by optimization or modifications due to constraints. This allows for the reader to more easily follow along with the mathematics involved, especially for highly specialized equations. The authors also take care to specify all of the thresholds that are used for filtering results. Thus the algorithm would be reproducible from the information provided. One could even go further to improve the system based upon the references and future work listed by the authors for features not include in their version of the program.

RATING

I would definitely accept this paper. There is solid justification for the proposed algorithm solving a relevant problem, the intuition for what the algorithm will be doing in each step is straightforward and primes the reader for the technical detail, and the implementation specifications are detailed enough for the result to be reproduced.

There are multiple applications for the face swapping. One application is to improve the aesthetics of services that handle images without the photographed individuals’ consent. One such service is Google Street View, which employs blurring to maintain privacy. A more uniform sharpness could be achieved by using face swapping of stock face images instead of blurring to protect identity. Another application is to composite images together, taking the best faces from each one. Combining burst photography with face swapping could result in a powerful tool for photography of human subjects. These are strong use cases for which this algorithm should be considered.

The details perform a face swap are well described and the results show the power of a well-designed algorithm. An optimal set of swaps can be chosen from over 30,000 images in under a second. When taking into account that this includes all of the transformation of the library images into the target image, this is a very efficient algorithm and the speed further justifies its practicality. Although a more qualitative test than a quantitative one, the results of the user study show that the face swapping algorithm does well in producing results that are indistinguishable from real life images. Over half of the images produced by the algorithm were marked as “real” from users in the blind study.

COMMENTS

Could this algorithm be extended into use for video as well? Similar to the blurring used on Google’s Street View, television media often must obfuscate faces of individuals. Naturally the face swap would not be able to reproduce the actions made by the target’s face, but there is potential for the swapped face to move with the head of target so they appear as one unit.