Institute for Homeland Security Solutions
Applied research • Focused results

June 2010

3D Scanning for Biometric Identification and Verification

Project Leads

Anselmo Lastra, PhD, University of North Carolina at Chapel Hill

Henry Fuchs, PhD, University of North Carolina at Chapel Hill

Greg Welch, PhD, University of North Carolina at Chapel Hill

Ali Farsaie, PhD, SIS, Inc.

Statement of Problem

Reliable and robust identification and verification of individuals is critical to homeland security applications such as surveillance, authorization for entry to secure areas, and passport identity verification. Traditional biometrics, such as mug shots, fingerprints, and voice recognition, have been used with some success. However, they exhibit serious disadvantages for some tasks. These three biometrics, for example, are problematic for surveillance (identification); even the traditional mug shot is difficult to use in automated surveillance applications because many factors, such as lighting and frontal visibility, cannot be controlled.

A relatively new biometric, 3D facial recognition, holds great promise. Although the technology is nascent, in a comprehensive 2006 study (Phillips et al., 2007) recognition performance using 3D shape and texture matched that of the much more mature technologies of high-resolution image recognition (which featured controlled lighting) and iris recognition. Additionally, 3D modeling promises to enhance recognition performance because it can be used to recognize people in profile as opposed to a typical forward-looking, mug-shot pose. Even when using 3D to match to a mug shot, an advantage is that a 3D model allows one to render a view of the person from any desired perspective—the pose, distance, and even lighting can be factored into the rendering to match any photos.

Scenarios in which 3D recognition could be profitably used include (a) verification of identity at an airport (for example the subject's face could be rapidly scanned while his or her smart-card ID is being examined, and the system could then match the scan with data on the ID); (b) identification at a secure site or even at an airport while people are walking down the hallways or standing in line; or (c) 3D pose extraction of a moving subject, thereby potentially enhancing recognition performance and enabling intent analysis.

This brief presents the technical background of the 3D scanning technologies, briefly surveys related biometrics that may be combined with 3D recognition, provides an overview of the major technical issues, and highlights research opportunities to overcome those issues.

Background

Probably the most studied technology for 3D modeling is baseline stereo vision. Two or more cameras image the scene, and corresponding points are selected in the images. If the cameras are calibrated (camera position and orientation, as well as lens and imager characteristics), the correspondences can be used via triangulation to determine the distance, and thus the geometry, of the visible structures in the scene.

A major problem is to determine the correspondences. A scene with very uniform color, such as white walls, is clearly problematic. If the scene is highly textured, however, then correspondence points may be extracted automatically. Stereo imagining techniques fall into two categories, intensity matching and feature detection (Faugeras, 1993; Forsyth & Ponce, 2002), with the latter having proven more reliable. Stereo reconstruction may also be performed from video sequences (Pollefeys & Gool, 2002; Pollefeys et al., 2004). The problem of accurately finding correspondences, however, has proven to be difficult and not always robust, leading researchers to investigate active approaches, primarily laser scanning and structured light.

Laser Scanning

When used for faces, human bodies, or other objects at short distances, triangulation is typical employed. A laser stripe scanned across the subject essentially provides correspondences for the camera(s). Cyberware makes a well known laser scanner of this type that has been extensively used in the movie industry. Unfortunately, this scanning technique takes anywhere from seconds to minutes—not a problem for scanning a seated and supported actor's face, but prohibitively long for identification purposes. Some laser techniques project complex patterns using interference of two beams, essentially the structured light technique described at the end of this section.

Another laser-scanning technique uses time of flight (the time for illumination to travel to and from a surface, divided by the speed of light) to determine distance. This is also known as LIDAR. Typically this method is used for longer ranges. Some new devices, such as the Swiss Ranger (from Mesa Imaging, http://www.mesa-imaging.ch) and Canesta (http://canesta.com) cameras work at ranges of a few meters and at video rates. However, their low resolution—160 x 120 pixels (Canesta) and 176 x 144 pixels (Swiss Ranger)—makes them unsuitable for biometrics. The marketing focus for these devices seems to be in vehicle safety applications (backup alarms for cars, for example) and human-computer interaction (potentially for video games).

Structured Light

The second general approach, *structured light*, is very similar to laser triangulation except that a light projector is typically used to project a pattern onto the subject. This provides a rich field of correspondences across the subject that can be used to extract a 3D model from the camera images. The use of time-multiplexed coded structured light patterns was first proposed by Posdamer and Altschuler (1982) and has sparked a great deal of research. Typically a small number of patterns are projected in sequence and the result imaged. Monochrome cameras can be used to capture geometry, and a color camera to add texture. This is the technology used by the 3D Snapshot system from SIS, Inc. The following sections focus on structured light (since it is the most suitable for human-subject scanning) and examine the challenges as well as possible research directions.

Issues

- The process should not disturb the subject. A major problem with conventional structured light approaches is that the rapidly flashing patterns are uncomfortable for the people being scanned. There may also be situations in which it would be important to scan a subject without his or her knowledge.
- Speed of capture is critical for any moving subject, especially for human biometrics. Many systems take less than a second (0.3 seconds for 3D Snapshot) to scan, but humans can move significantly in that time. An ideal scan duration would be from 1/10th to 1/30th of a second.
- Speed of processing is also important. The result must be available within a second or two. Ideally, the processing could be done at real-time rates in order to generate 3D at video rates.
- Accuracy is a major issue, of course, especially under less-than-ideal lighting and environmental conditions.
- The scanner should have a reasonably wide field of view so the subject does not have to be in a very precise location. Analogously, the scanning device should have reasonable depth of field.

- Eyeglasses are a problem because of reflections from, and refraction through, the lenses.
- Geometry of hair can be difficult to capture, and a beard can also be used to hide features.

Research Directions

Research directions in this section are proposed in priority order, based on the importance of the problem to be solved, as well as the amount of time expected to develop a technical solution.

Imperceptible Scanning

The authors see two fruitful technical directions to make the scanning process invisible to the subject. The first, *imperceptible structured light*, was invented at the University of North Carolina at Chapel Hill (Raskar et al., 1998) to enable 3D modeling of persons for 3D video conferencing applications. The key idea of imperceptible structured light is to flash a pattern and its inverse rapidly enough that it will appear to the subject as white light. A fast camera can be synchronized to the projector and will capture an image of the pattern. Most of the work in this area has been to calibrate projector systems shining on non-planar environments (Cotting, Naef, Gross, & Fuchs, 2004; Cotting, Ziegler, Gross, & Fuchs, 2005; Zollmann & Bimber, 2007). Although the authors have demonstrated the concept, many challenges remain with the hardware implementation.

The other potential approach is to use infrared illumination. Infrared may be imaged directly (essentially to detect skin temperature) (Abayowa, 2009; Colantonio & Benvenuti, 2007), or infrared patterns can be projected, much as with visible light. There has been little work on infrared structured light. The authors know only of a bench prototype tested in Japan (Akasak, Sagawa, & Yagi, 2007).

Speed

Two factors account for the time required for a scan: acquisition and processing. Carefully synchronizing the camera with the projector, such as the authors have done with their prototypes (Cotting et al., 2004; Cotting et al., 2005; Raskar et al., 1998), can make the image acquisition process faster. However, imaging in a shorter amount of time, or with less light, tends to make sensor noise more problematic, and this should be combated using techniques such as those of Bennett and McMillan (2005). To make the processing faster, the authors can use the graphics processing unit, an approach they pioneered (Harris, Coombe, Scheuermann, & Lastra, 2002) that is now becoming popular. Speed increases of 20 to 40 times are possible.

Improved Biometric Accuracy

It is possible to combine multiple biometrics, with the resulting biometric fusion potentially increasing accuracy. A promising approach may be to combine iris/retinal scanning with 3D scanning. The texture of the human iris forms during the gestational period, and it exhibits a great deal of detail, including furrows, freckles, and other features (Daugman, 2004). The iris can be imaged unobtrusively, and the near-infrared modality used brings out patterns even in persons with dark pigmentation. Because imaging of the iris requires cooperation from the subject, however, it may be less useful for identification from surveillance imagery (Abayowa, 2009). A survey of techniques is presented in Bowyer, Hollingsworth, & Flynn (2008).

Field of View and Depth of Field

The ability to capture 3D models of people over a wide working area will provide a very powerful biometric tool. This is a very difficult problem, however. For the hardware part of the solution, the authors propose overlapping, synchronized structured light projectors and a set of cameras. The prices are dropping rapidly for both of these devices, so cost is not the primary barrier.

This net of projectors and cameras could be coupled with software algorithms for a progressive refinement of the biometric over time. For example, the scanning might occur as people are standing at the line waiting for the TSA screening. Even if there is no wait, just the walk through the cordoned area could serve.

A potentially powerful strategy is to combine structured light approaches with extraction of correspondences for a combined modeling approach. The longer observation time allowed in the screening-while-walking scenario can be used to improve the models by predicting the subject's motion and tailoring the imperceptible structured-light patterns to improve the model.

Extraction of Subject Pose and Posture

The way a person walks is a very characteristic identifier for recognizing someone. Furthermore, pose and posture analysis could be used to analyze intent in certain situations. The authors have been working with the Navy to estimate the posture of Marines during training and using the posture to analyze their performance. Because a multitude of views is necessary, this work is using multiple video cameras. Structured light would be a very useful enhancement that is not possible for the outdoor Marine-training scenario.

Contact Information

Anselmo Lastra
Sitterson Hall, CB3175, University of North Carolina at Chapel Hill
Chapel Hill, NC 27599-3175
919-962-1958
lastra@cs.unc.edu

Anselmo Lastra, PhD, is a Professor and Chairman of Computer Science at the University of North Carolina at Chapel Hill. He received a BS in electrical engineering from the Georgia Institute of Technology and MS and PhD degrees in computer science from Duke University. He was co-chair of I3D 2005 and Graphics Hardware 2000 and 2004 and is an associate editor of *IEEE Computer Graphics and Applications*. His research interests are in the areas of image-based 3D modeling and rendering and graphics hardware architectures.

Henry Fuchs, PhD, is the Federico Gil Professor of Computer Science at the University of North Carolina at Chapel Hill. He has been active in computer graphics since the early 1970s, with rendering algorithms (BSP Trees), hardware (Pixel-Planes and PixelFlow), virtual environments, tele-immersion systems, and medical applications. He is a member of the National Academy of Engineering, a fellow of the American Academy of Arts and Sciences, and the recipient of the 1992 ACM-SIGGRAPH Achievement Award, the 1992 Academic Award of the National Computer Graphics Association, and the 1997 Satava Award of the Medicine Meets Virtual Reality Conference.

Greg Welch, PhD, is a research associate professor of computer science at the University of North Carolina at Chapel Hill. In 1986 he received a degree in Electrical Technology from Purdue University (with Highest Distinction), and in 1996 he received a PhD in computer science from the University of North Carolina at Chapel Hill. Previously he has worked at NASA's Jet Propulsion Laboratory and Northrop-Grumman's Defense Systems Division. His research interests include human-tracking systems, 3D telepresence, projector-based graphics, computer vision and view synthesis, and medical applications of computers. He has co-authored over 50 peer-reviewed publications in these areas, is a co-inventor on multiple patents, and maintains an internationally recognized Web site dedicated to the Kalman filter. He is a member of the IEEE Computer Society and ACM.

Ali Farsaie, PhD, is president and CEO of Spatial Integrated Systems, Inc. (SIS). He provides strategic guidance and development for new program activities within defense, aerospace, Federal agencies, and commercial activities. Governor Perdue recently appointed him to North Carolina’s Economic Development Board. Dr. Farsaie received MS and PhD degrees in 1979 from North Carolina State University. Prior to SIS, Dr Farsaie was a chief engineer at the Naval Surface Warfare Center Dahlgren Division. He conducted, formulated, and managed research and development of novel approaches in advanced information technology, sensors and robotics, training, and system engineering to meet long-term Navy mission requirements.

References

Abayowa, B. O. (2009). Thermal infrared exploitation for 3D face reconstruction. *Proceedings of SPIE - The International Society for Optical Engineering, SPIE-7347*.

Akasak, K., Sagawa, R., & Yagi, Y. (2007). *A sensor for simultaneously capturing texture and shape by projecting structured infrared light.* Paper presented at the 3DIM.

Bennett, E. P., & McMillan, L. (2005). *Video enhancement using per-pixel virtual exposures*. Paper presented at the ACM SIGGRAPH.

Bowyer, K. W., Hollingsworth, K., & Flynn, P. J. (2008). Image understanding for iris biometrics: A survey. *Computer Vision and Image Understanding, 110*(2), 281-307.

Colantonio, S., & Benvenuti, M. (2007). Object tracking in a stereo and infrared vision system. *Infrared Physics & Technology, 49*(3), 266-271.

Cotting, D., Naef, M., Gross, M., & Fuchs, H. (2004). *Embedding imperceptible patterns into projected images for simultaneous acquisition and display.* Paper presented at the ISMAR.

Cotting, D., Ziegler, R., Gross, M., & Fuchs, H. (2005). *Adaptive instant displays: Continuously calibrated projections using per-pixel light control.* Paper presented at Eurographics.

Daugman, J. (2004). How iris recognition works. *IEEE Transactions on Circuits and Systems for Video Technology, 14*(1), 21-30.

Faugeras, O. (1993). *Three-dimensional computer vision: A geometric viewpoint*. Cambridge: The MIT Press.

Forsyth, D. A., & Ponce, J. (2002). *Computer vision: A modern approach*: Prentice Hall.

Harris, M. J., Coombe, G., Scheuermann, T., & Lastra, A. (2002). *Physically-based visual simulation on graphics hardware.* Paper presented at the SIGGRAPH / Eurographics Workshop on Graphics Hardware.

Phillips, P. J., Scruggs, W. T., O'Toole, A. J., Flynn, P. J., Bowyer, K. W., Schott, C. L., et al. (2007). *FRVT 2006 and ICE 2006 large-scale results* (No. NISTIR 7408). Gaithersburg, MD: National Institute of Standards and Technology.

Pollefeys, M., & Gool, L. V. (2002). From images to 3D models. *Communications of the ACM, 45*(7), 50-55.

Pollefeys, M., Van Gool, L., Vergauwen, M., Verbiest, F., Cornelis, K., Tops, J., et al. (2004). Visual modeling with a hand-held camera. *International Journal of Computer Vision, 59*(3), 207-232.

Posdamer, J. L., & Altschuler, M. D. (1982). Surface measurement by space-encoded projected beam systems. *Computer Graphics and Image Processing*.

Raskar, R., Welch, G., Cutts, M., Lake, A., Stesin, L., & Fuchs, H. (1998). *The office of the future: A unified approach to image-based modeling and spatially immersive displays.* Paper presented at the SIGGRAPH98, Orlando, FL.

Zollmann, S., & Bimber, O. (2007). *Imperceptible calibration for radiometric compensation.* Paper presented at Eurographics.