Multi-view Image Processing: Current State of the Art, Challenges, and Future Directions

Dapeng Oliver Wu
Electrical & Computer Engineering
University of Florida
August 18, 2006
http://www.wu.ece.ufl.edu/

Outline

- Definition of multiple views
- Basic multi-view image processing techniques
- Applications of multi-view image processing
- Grand challenges and future directions

Definition of Multiple Views

- Multiple views from different positions or angles (space)
  - E.g., images captured by multiple cameras at the same time
- Multiple views from different time instants
  - E.g., satellite images captured at different times
- Multiple views from different imaging modalities
  - E.g., CT, MRI, Acoustic images

Outline

- Definition of multiple views
- Basic multi-view image processing techniques
- Applications of multi-view image processing
- Grand challenges and future directions

Image Enhancement

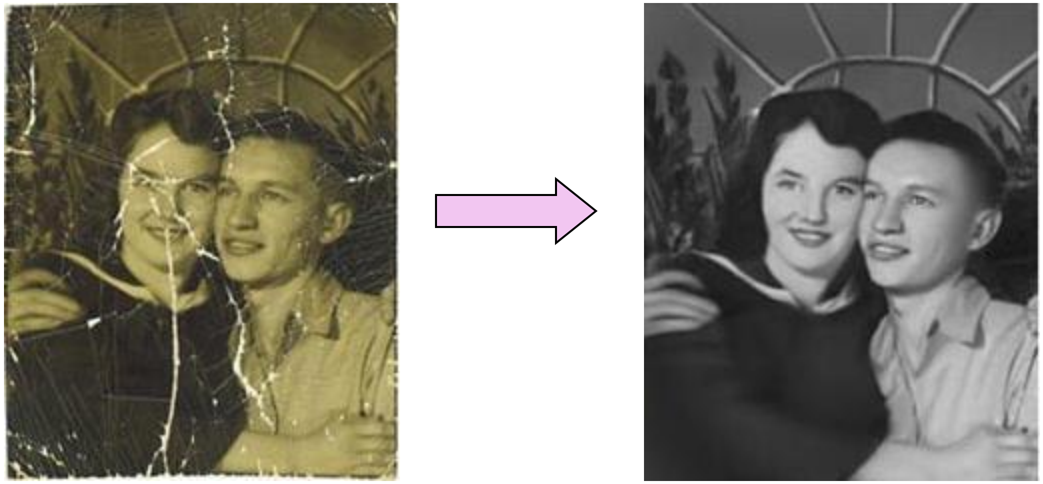
- Multi-view improves estimation of illumination ⇨ mitigate variations in illumination, provides better image quality

Bennett Wilburn et al., "High performance imaging using large camera arrays," ACM SIGGRAPH 2005

Image Restoration

- Multi-view improves estimation of image degradation function, parameter estimation

  ⇨ better image restoration



Super-resolution

- Multi-view provides space diversity

⇨ improve spatial resolution

S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Advances and Challenges in Super-Resolution", International Journal of Imaging Systems and Technology, Special Issue on High Resolution Image Reconstruction, vol. 14, no. 2, pp. 47-57, August 2004

Digital Inpainting

- It is image interpolation in a general sense
- First introduced in the paper by Bertalmio-Sapiro-Caselles-Ballester (SIGGRAPH 2000)
- Broad applications
  - image interpolation
  - photo restoration
  - zooming
  - super-resolution
  - primal-sketch based perceptual image compression
  - error concealment of (wireless) image transmission
- Multiview improves estimation of digital inpainting model parameters

Image Coding

- Multi-view provides spatial correlation

  ⇨ more efficient compression

- Multi-view image/video coding
  - Motion-compensation over time and view; one I-frame is enough for N views; save (N-1) I-frames
  - Can we just code 3D images + view angles instead of 2D images? (future direction)

Image Segmentation

- Image segmentation
  - Multi-view provides more information about the scene and objects ⇨ better segmentation

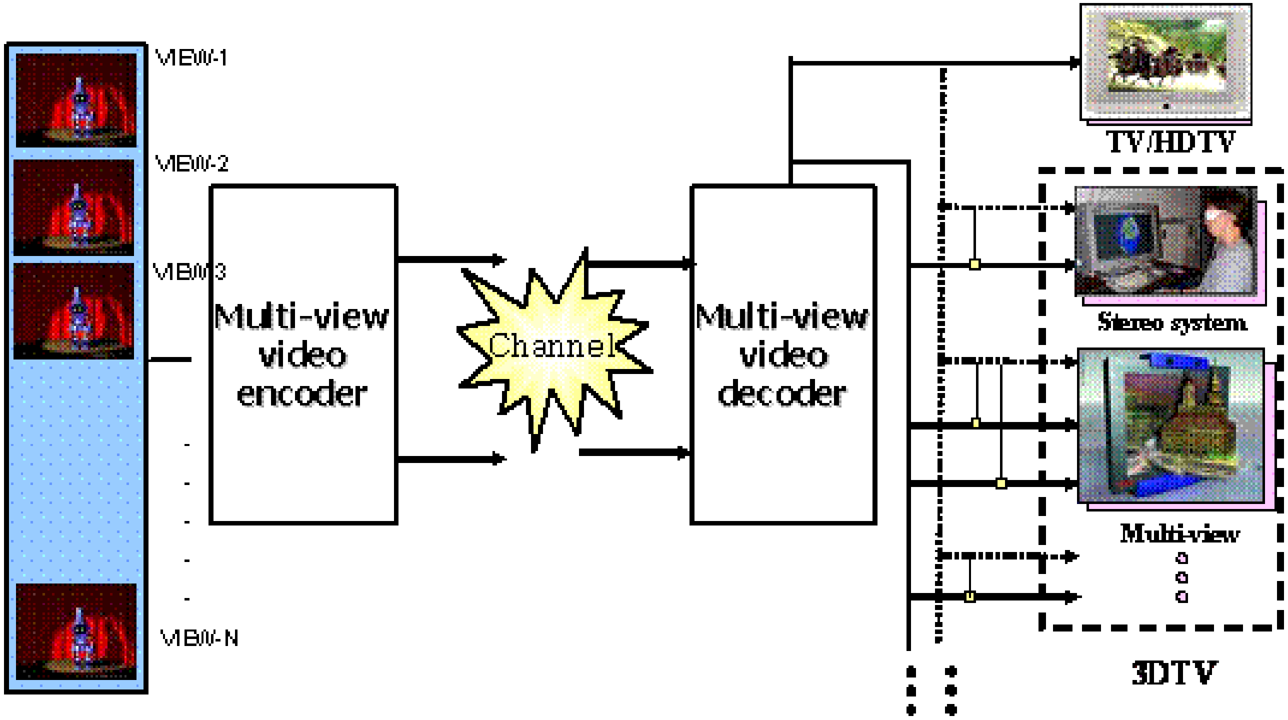
Object Recognition

- Multi-view provides more information about an object ⇒ higher recognition probability
- E.g., The story of six blind men and an elephant

Outline

- Definition of multiple views
- Basic multi-view image processing techniques
- Applications of multi-view image processing
- Grand challenges and future directions

Multi-view Video System



"Introduction to Multi-view Video Coding," ISO/IEC JTC 1/SC 29/WG 11N7328, July 2005, Poznan, Poland

Applications of Multi-view Video

- Free viewpoint video, free viewpoint TV, 3D video, 3D-TV
- Entertainment – concert, sports, multi-user game, movie, drama, news
- Education – cultural archives, manual with real video, instruction of sports playing, medical surgery
- Sightseeing – zoo, aquarium, botanical garden, museum.
- Surveillance – traffic intersection, underground parking, bank
- Archive – space archive, living national treasures, traditional entertainment
- Art/Content – creation of new type of media art and digital content
- Production – special effects
- Broadcasting

3D Video Display

- Holography
- 3D monitor with special glasses (red/green, polarization)
- 3D monitor without glasses (Sharp, SeeReal)
- 3D wide screen cinema with special glasses

Immersive TV

- E.g., use virtual reality techniques

Figure 1: Immersive Television using Head mounted displays (left) or wide screen projection (right)

R. Schäfer and P. Kauff and O. Schreer, "Immersive Telepresence for Broadcast and Communication," Proc. of IST 2000, Nice, France, November 2000.

Immersive Teleconference

Figure 2: Long term vision (left) and medium term realisation (right) of immersive teleconference systems

R. Schäfer and P. Kauff and O. Schreer, “Immersive Telepresence for Broadcast and Communication,” Proc. of IST 2000, Nice, France, November 2000.

Interactive Free-Viewpoint Video

- Free-viewpoint video
  - allows rendering of real events from novel views
  - allows user interactive control of viewpoint with a visual-quality comparable to captured video
- Applications
  - Broadcast production: freeze-time visual effect
  - Video games
  - Visual communication
- Techniques:
  - Visual hull
  - Photo hull
  - Image-based rendering, segmentation, image correspondence, volumetric geometry

Free-viewpoint Television (FVT)

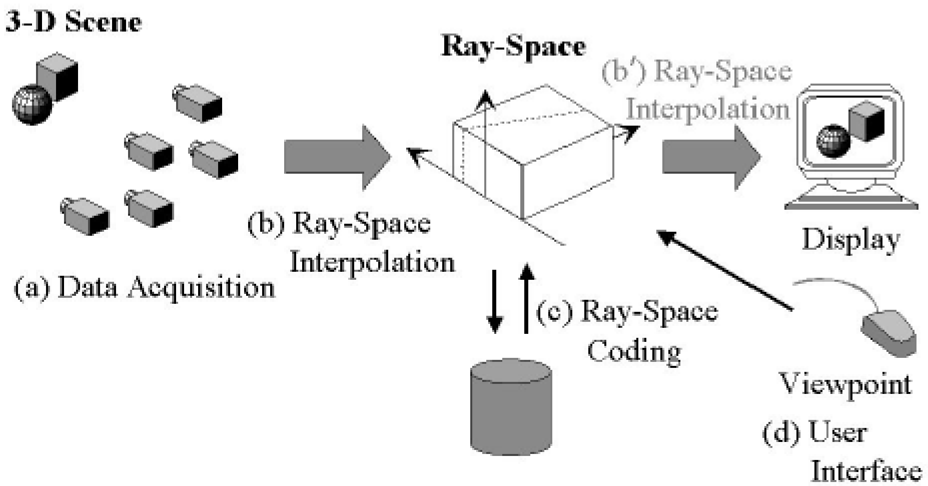


Figure 1: Configuration of FTV system.

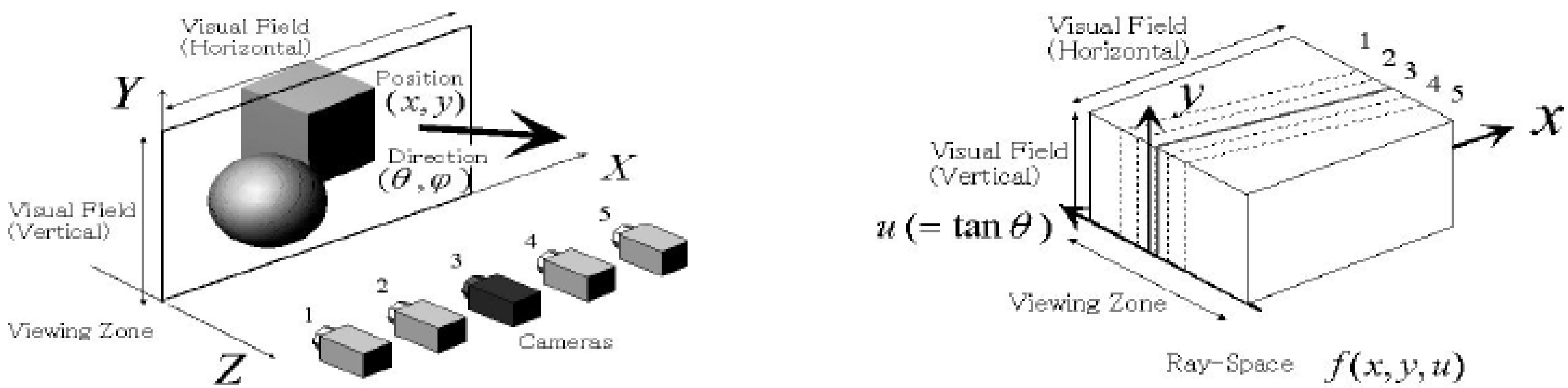
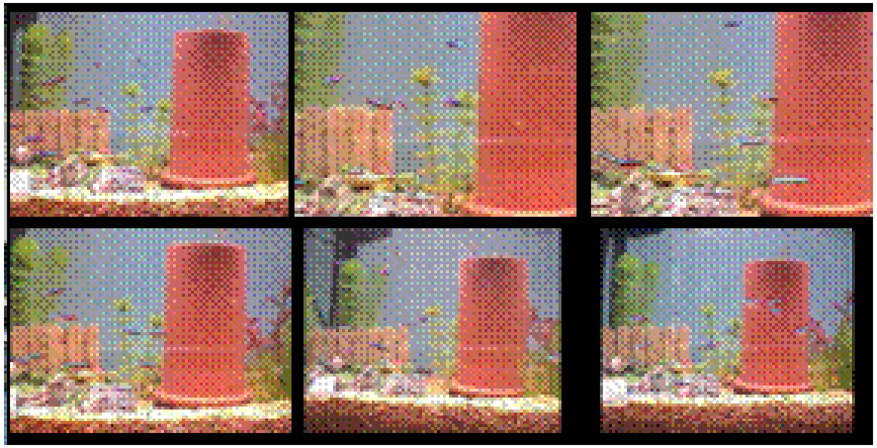


Figure 2: Acquisition of FTV signal.

Example of FVT Images

- Pan
- Zoom in, zoom out



Masayuki Tanimoto, “Free Viewpoint Television - FTV”, Proc. of Picture Coding Symposium, San Francisco, CA, USA, December 15-17, 2004.

Sports Video Content Analysis and Application

- Multiview helps measurements in sports
  - Velocity
  - Force
  - Angle
  - Position
  - Alignment of various parts of body
  - Trajectory of balls

High Performance Imaging Using Large Camera Arrays

- Image mosaic: stitch multiple images into a panoramic view
- High speed video capture: convert space into time
  - E.g., 52 cameras, each with 30 fps, provide 1560 fps
  - Slowly replay high speed event (e.g., a moving bullet through balloons)

Synthetic Aperture Photography

(a) (b) (c)

Figure 11: Matted synthetic aperture photography. (a) A sample image from one of 90 cameras used for this experiment. (b) The synthetic aperture image focused on the plane of the people, computed by aligning and averaging images from all 90 cameras as described in the text. (c) Suppressing contributions from static pixels in each camera yields a more vivid view of the scene behind the occluder. The person and stuffed toy are more clearly visible.

Bennett Wilburn et al., “High performance imaging using large camera arrays,” ACM SIGGRAPH 2005

3D Building Reconstruction from Aerial Imagery

- LIDAR (Light Detection and Ranging; or Laser Imaging Detection and Ranging) is a technology that determines distance to an object or surface using laser pulses.

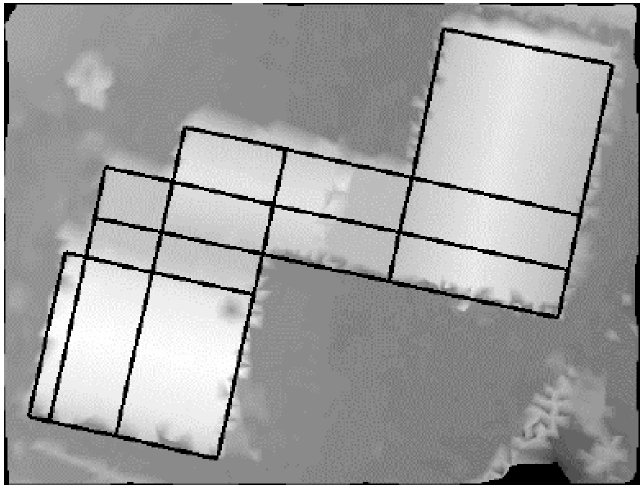


Figure 2: Partitioned building outline as overlay on grey value coded heights.

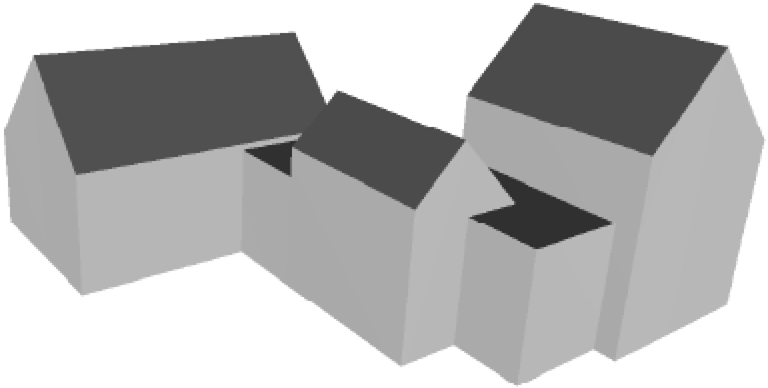


Figure 9: Reconstructed 3D building model.

G. Vosselman and S.T. Dijkman, "3D building model reconstruction from point clouds and ground plans," International Archives of Photogrammetry and Remote Sensing, vol. XXXIV, part 3/W4, pp. 37-44, 2001.

3D Building Reconstruction from Aerial Imagery (2)

- 3D building models become increasingly popular among urban planners and the telecommunication industry.
- Analysis of propagation of noise and air pollution through cities and estimation of real estate taxes are some other potential applications of 3D building models.

Multi-modality Image Processing: Image Fusion

- Multiple modalities: e.g., to detect breast cancer, use
  1. Planar X-ray imaging
  2. X-ray computed tomography (CT)
  3. Magnetic resonance imaging (MRI)
  4. Planar scintigraphy
  5. Single photon emission computed tomography (SPECT)
  6. Positron emission tomography (PET)
  7. Ultrasonic imaging
  8. Thermoacoustic tomography (TAT)
  9. RF-induced thermoacoustic tomography
  10. Photoacoustic tomography (PAT)
      - Laser-induced photoacoustic tomography
  11. Optical tomography
      - Ultrasound-modulated optical tomography (UOT)
  12. Optical-coherence tomography (OCT)
      - Mueller optical-coherence tomography

Outline

- Definition of multiple views
- Basic multi-view image processing techniques
- Applications of multi-view image processing
- Grand challenges and future directions

Grand Challenges

- Cosmic scale
  - Satellite imagery
  - Radio telescope imagery
- Conventional scale
- Micro/nano scale
  - The best resolution for an optical microscope is about 0.2 microns = 200 nm, limited by visible light diffraction.
  - Cells have size of 10's microns.

Grand Challenge: Cosmic Scale

- Can we improve the resolution of current satellite imagery by a factor of 10 or even 100?
- Potential impact: why this is important?
  - Able to see details never available previously
  - Particularly important for intelligence, Department of Defense, and homeland security
- Possible solutions
  - Multi-view image processing of multiple satellite images

Grand Challenge: Micro/Nano Scale

- Can we improve the resolution of current optical microscope by a factor of 100, i.e., 2 nm?
- Potential impact: why this is important?
  - It substantially breaks the physical limit! Nobel prize level research!
  - Able to see nano-scale activities in live organisms never available previously, e.g., see live DNA molecules (the diameter of the DNA helix is 2 nm)
  - Particularly important for physiology and medicine (e.g., gene therapy)
  - Significantly improve current micro-array technologies
- Possible solutions
  - Multi-view image processing of multiple microscope images

Computed Tomography (CT)

- Allan M. Cormack and Godfrey N. Hounsfield were awarded Nobel Prize in Physiology or Medicine in 1979 for development of CT
- How CT was invented?
  - Cormack proposed the back-projection method to generate a three-dimensional image of the internals of a human body from a large series of two-dimensional X-ray images taken around a single axis of rotation.
  - Hounsfield constructed the first CT system practicable in medical care.

Magnetic Resonance Imaging (MRI)

- Paul C Lauterbur and Peter Mansfield were awarded Nobel Prize in Physiology or Medicine in 2003 for development of MRI
- How MRI was invented?
  - Paul Lauterbur proposed a method that uses gradient magnets along with the main magnet to detect signals for creating 2D images.
  - Peter Mansfield proposed a signal processing method to transform the detected signals into a 2D image.

Several Nobel Prizes Related to MRI

- Felix Bloch and Edward Mills Purcell were awarded the Nobel Prize in Physics in 1952 for their discovery of nuclear magnetic resonance (NMR).
- Richard Ernst was awarded the Nobel Prize in Chemistry in 1991 for his contributions to the development of the methodology of high resolution NMR spectroscopy.
- Kurt Wüthrich was awarded the Nobel Prize in Chemistry in 2002 for his development of NMR spectroscopy for determination of the three-dimensional structure of biological macromolecules in solution.

Another Nobel Prize on Imaging

- Dennis Gabor was awarded Nobel Prize in Physics in 1971 for his invention and development of the holographic method
- How holograph was invented?
  - ?

Grand Challenge: Conventional Scale

- Convert single 2D video sequence to 3D video sequence?
- Super-resolution?
- …

Future Directions

- Multi-view video coding:
  - code 3D images + view angles instead of 2D images

Computer-assisted Surgery

- 3D CT/MRI images can help locate tumors to be removed.
- But an imprecise cut will cause severe damage to patients, which sometimes is irreversible.
- Computer-assisted surgery can help surgeons to determine the precise location of tumors and do practice before the surgery.
- Computer-assisted surgery can also provide risk assessment for patients.

Future Directions (cont’d)

- The list goes on and on.

Thank you!