

# Stereoscopic 3D (TV) 영상처리

연구개발 동향 및 기술개요

오윤택, ph.D.

KJIST U-VR Lab  
<http://vr.kjist.ac.kr>

## Outline

- ◆ Introduction
  - ◆ Motivation and research activities
- ◆ Multiview Image Acquisition
  - ◆ Camera geometry and 3D distortion
- ◆ Stereo Video Compression
  - ◆ Scalability tools and Multiview profile
- ◆ Functionality for 3D Applications
  - ◆ View synthesis & pi-VE
- ◆ Summary and Conclusion

## Introduction: KJIST U-VR Lab

### ◆ Main focus of *U-VR Lab*

- ◆ PUI for smart environment in 10 years
- ◆ 3D 실감통신: 가상 vs. 현실 (VR/AR/MR)

### ◆ What's PUI?

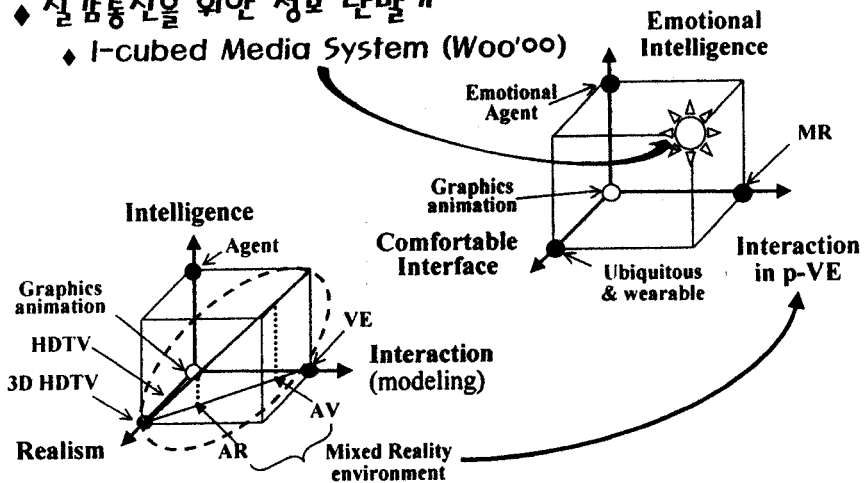
- ◆ 개인화 (Personalization)
  - ◆ 지능화 (perceptual/emotional intelligence)
- ◆ 편재성 (Ubiquity)
  - ◆ 편재된 컴퓨팅 (ubiquitous/wearable computing)
- ◆ 통합화 (Internetworking)
  - ◆ 유무선 통신: 광(고속) vs. 무선(개인화)



## Introduction: 실감통신 요소기술

### ◆ 실감통신을 위한 정보 단말기

#### ◆ I-cubed Media System (Woo'00)



## Introduction: 실감통신기술 발전과정

### ◆ Audio-visual Communications

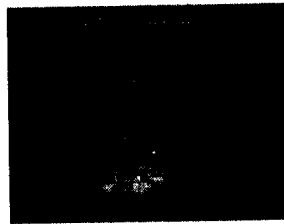
- ◆ 음성통신: Mono-> Stereo -> 3D sound
- ◆ 영상통신: (motion+color+HD+?)
  - ◆ 사진->흑백TV->컬러TV->HDTV(1980)->?

### ◆ What's next?

- ◆ 오감통신: 시각, 청각, 후각, 미각, 촉각
- ◆ 육감통신? emotion & imagination

### ◆ Where is 3D(TV)?

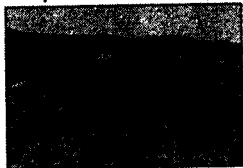
- ◆ 실감통신을 위한 정보 단말기
- ◆ 3D HDTV vs. 3D HMD



## Introduction: Why 3차원 영상?

### ◆ Why 3D with multiview?

- ◆ 실감통신: 임장감, 실존감, 자연감
  - ◆ Mono 3D cues: Focus length, linear/aerial perspective, texture gradient, light & shade, occlusion, motion parallax, etc.
  - ◆ 3D with Stereo-cues: binocular disparity (Human: 65mm), convergence angle, etc.
- ◆ (Object-based) functionality (and interactivity)
- ◆ 응용분야: 정보통신, 방송, 의료, 교육/훈련, 군사, animation, Game, VR/MR/AR, CAD, etc.

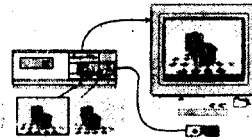


## Introduction: 3차원 영상 단말기

### ◆ What about 3D Display?

#### ◆ 안경식: 편광방식과 시분할방식

- ◆ Free-view
- ◆ Anaglyph
- ◆ Polarized glasses
- ◆ Shutter glasses



#### ◆ 비안경식

- ◆ Volumetric display
- ◆ Stereoscopic : Barrier vs. Lenticular



#### ◆ What's next? 지능형 3D 단말기

- ◆ 3D HDTV vs. Hologram (volumetric)
- ◆ See-through 3D HMD (AR)



## Introduction: 3차원 영상처리 연구동향

### ◆ 유럽

- ◆ Cost230: 3DTV
- ◆ DISTIMA (92-95): 3DTV
- ◆ PANORAMA (96-98): 3D Telepresence



### ◆ 일본:

- ◆ TAO, NHK, NTT, SANYO, ATR

### ◆ 미국:

- ◆ CMU, MIT, NASA-JPL, USC, SRI, etc.

### ◆ 한국

- ◆ 원자력연구소, KIST, SAIT
- ◆ KAIST, KJIST, KNU, POSTECH, 광운대 etc.



## Introduction: 3차원 영상처리 연구동향

### ◆ RACE2405: DISTIMA (92-95)

- ◆ To develop SW/HW to realize 3D TV
- ◆ Approach: MPEG-2 Compatible Codec
  - ◆ Video over ATM (9Mbps): left (6-7Mbps), right (3-2Mbps)
  - ◆ Up to 4 video channels w/ Dolby surround sound



### ◆ ACT5 AC044: MIRAGE (-98)

- ◆ virtual studio



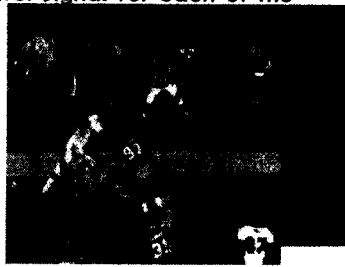
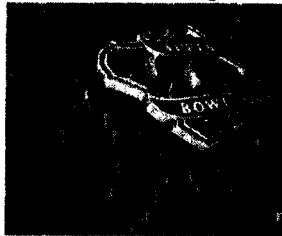
### ◆ ACT5-AC092: PANORAMA (95-98)

- ◆ To develop SW/HW to realize 3D telepresence
- ◆ Approach
  - ◆ WPG1: synthesis (parametric 3-D scene description)
  - ◆ WPG2: analysis (coding & intermediate-view synthesis)
  - ◆ WPG3: 3-cam's, autostereoscopic display, headtracking

## Introduction: 3차원 영상처리 연구동향

### ◆ EyeVision: 3D goes to Superbowl (CBS & CMU)

- ◆ Super Bowl XXXV @ Raymond James Stadium, FL
- ◆ How does the system work?
  - ◆ about 30 cameras w/ computer-controlled zoom and focus
  - ◆ A human operator manipulates a movable pan-tilt tripod
    - > The tripod is equipped with sensors to measure its angle
  - ◆ The master cam-head mimicks the motion of the tripod
  - ◆ Information from the master cam is fed to a computer
    - > pan-tilt angles, zoom, focus, etc.
  - ◆ The computer computes the control signal for each of the remaining cameras

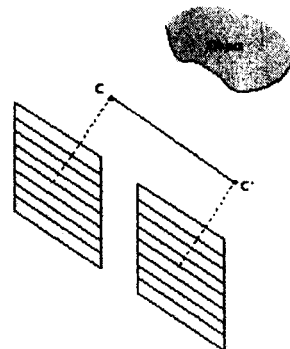
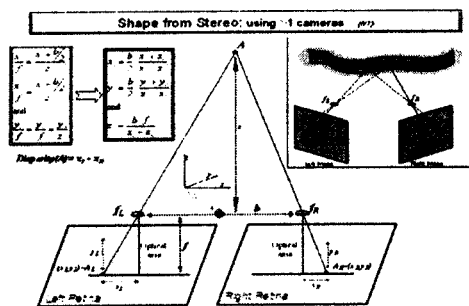


## Outline

- ◆ Introduction
  - ◆ Motivation and research activities
- ◆ Multiview Image Acquisition
  - ◆ Camera Geometry and 3D Distortion
- ◆ Stereo Video Compression
  - ◆ Scalability tools and Multiview Profile
- ◆ Functionalities for 3D Applications
- ◆ Summary and Conclusion

## 3차원 영상 획득

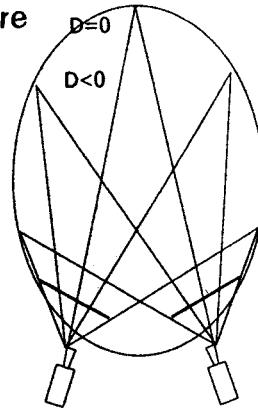
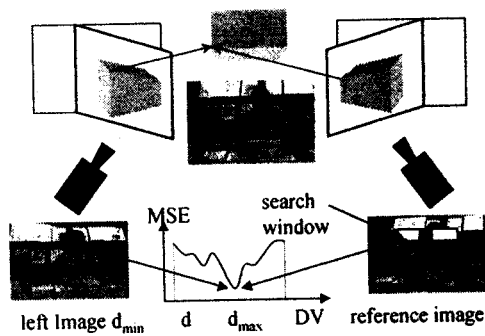
- ◆ 3D image acquisition
- ◆ Camera Setup: Parallel
  - ◆ Simple DE: horizontal direction disparity
  - ◆ No 3D distortion: depth plane is linear



## 3차원 영상회전

### ◆ Camera Setup: Angled Camera Setup

- ◆ Keystoning: vertical direction error  $D > 0$
- ◆ Eye fatigue: Depth plane curvature  $D = 0$
- ◆ requires adjustment before DE  $D < 0$

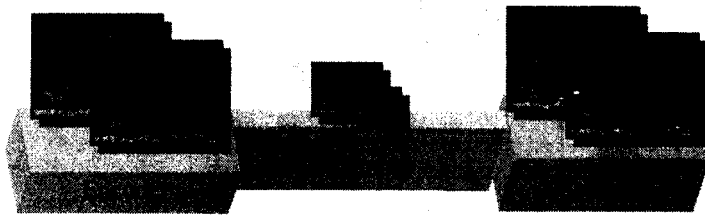


## Outline

- ◆ Introduction
  - ◆ Motivation and research activities
- ◆ Multiview Image Acquisition
  - ◆ Camera Geometry and 3D Distortion
- ◆ Multiview Video Compression
  - ◆ Scalability tools and Multiview Profile
- ◆ Functionalities for 3D Applications
- ◆ Summary and Conclusion

## 3차원 영상압축

- ◆ Why 3D with multiview?
  - ◆ 실감통신: 임장감, 실존감, 자연감
  - ◆ (Object-based) functionality (interactivity)
- ◆ Trade-offs: 실감 vs. 대정보량
  - ◆ Channel BW, protocols, data amount



## 3차원 영상압축

- ◆ Main Coding Issues
  - ◆ Coding Efficiency
    - ◆ Occlusion detection and treatment
    - ◆ Joint motion/disparity estimation
  - ◆ Compatibility with Standards
    - ◆ MPEG-2: scalability
    - ◆ MPEG-4: object-based scalability
  - ◆ Functionality (interactivity)
    - ◆ Accurate & smooth disparity estimation
    - ◆ Segmentation or object-based coding

## 3차원 영상압축

### ◆ Problem Formulation: (Joint Optimization)

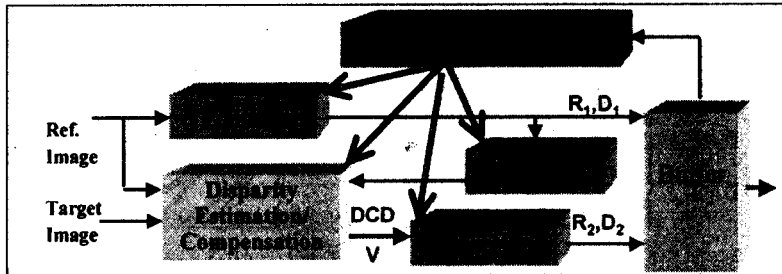
Given  $F_1, F_2, R_{budget}$

Find  $\hat{X} = (V, Q_1, Q_2)$

such that  $\hat{X} = \arg \min_X \{D_1(Q_1) + \alpha \cdot D_2(Q_2 | Q_1, V)\}$

subject to  $R_1(Q_1) + R_2(Q_2 | Q_1, V) \leq R_{budget}$

where "α" supports *Fusion vs. Suppression theory*

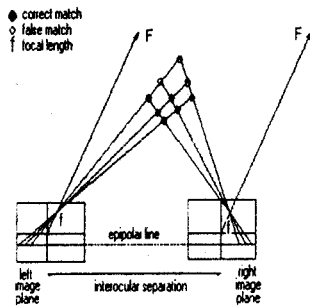


## 3차원 영상압축

### ◆ Motivation of 3D Coding

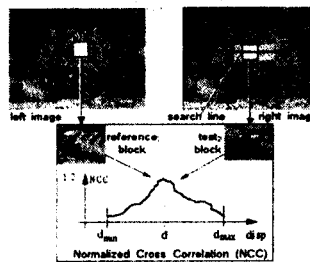
- ◆ Bottleneck: limited channel bandwidth
- ◆ Redundancy: temporal & binocular

### ◆ Disparity Estimation



Disparity estimation using block-matching

Stereo-camera setup -> one-dimensional search

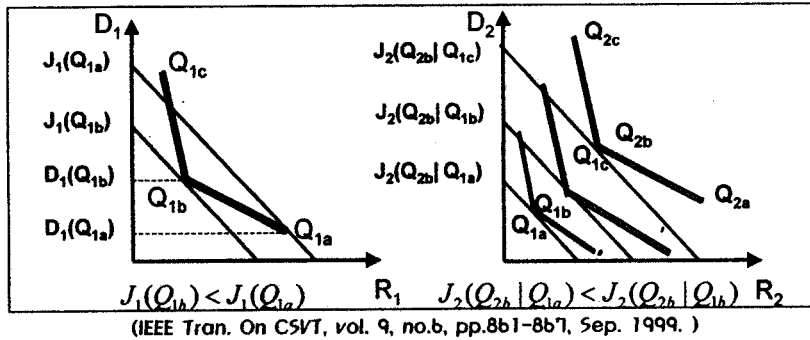


### 3차원 영상압축: Main Coding Issues

◆ Coding Efficiency

◆ Blockwise Dependent Quantization

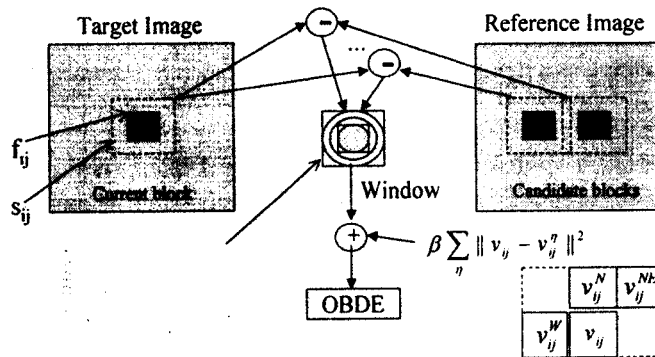
- ◆ ORD plot with 3 different quantizers
  - > Lagrangian cost:  $J = J_1 + J_2$ , where  $J_i = D_i + \lambda R_i$ ,  $i=1,2$
  - > can be  $J_1(Q_{1a}) + J_2(Q_{2b}|Q_{1a}) < J_1(Q_{1b}) + J_2(Q_{2b}|Q_{1b})$



### 3차원 영상압축: Main Coding Issues

◆ Coding Efficiency

◆ DE w/ FSBM & DC w/ Adaptive OBM



$$DPCM(v_{ij}) = v_{ij} - \text{median}(v_{ij}^W, v_{ij}^N, v_{ij}^{NE})$$

(IEEE Tran. On CSVT, vol. 9, no.6, pp. 194-200, Mar. 2000)

### **3차원 영상압축: Scalability**

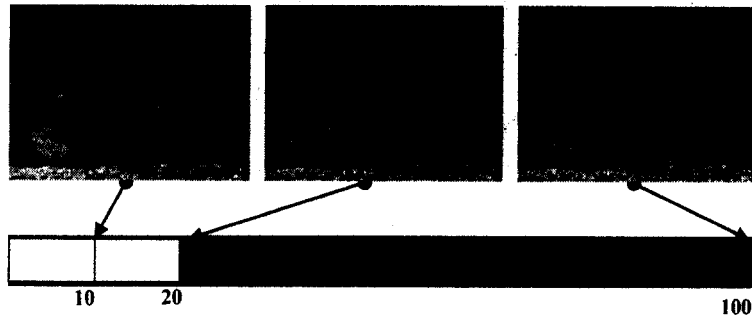
- ◆ Why scalability?
  - ◆ Priority: error resilience on noisy channel
  - ◆ Multi-quality video services (VOD, HDTV, etc.)
  - ◆ Internetworking of standards or equipment
- ◆ Basic Idea of Scalable Coding
  - ◆ Layered or hierarchical coding
  - ◆ Independent coding of the lowest layer
  - ◆ Dependent coding of each following layers
  - ◆ Coding complexity & quality scalability

### **3차원 영상압축: Scalability Tools**

- ◆ Data Partitioning
  - ◆ Break a coded bit-stream into essential & additional parts
- ◆ SNR (Quantization Noise) Scalability
- ◆ Spatial (Resolution) Scalability
- ◆ Temporal (Resolution) Scalability
  
- ◆ Hybrid Scalability

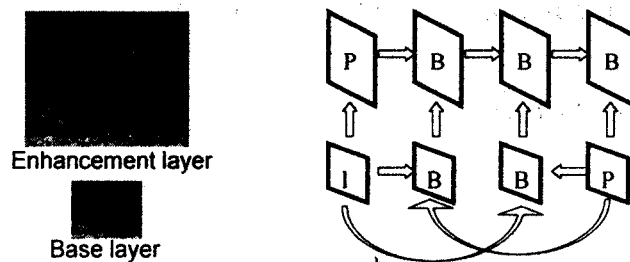
### 3차원 영상압축: SNR Scalability

- ◆ Control Quantization Step
  - ◆ Quantization Noise Scalability
  - ◆ Each layer coded at the same resolution w/ different quality



### 3차원 영상압축: Spatial Scalability

- ◆ Extended Pyramid Coding
  - ◆ Base layer: coded at lower (sampling) resolution
  - ◆ Enhancement layer: upsampled and predicted from the BL
- ◆  Backward compatibility: H.26x, MPEG-1.

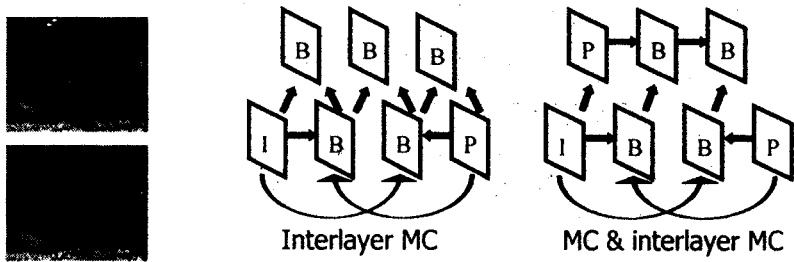


### 3차원 영상압축: Temporal Scalability

◆ Basic idea

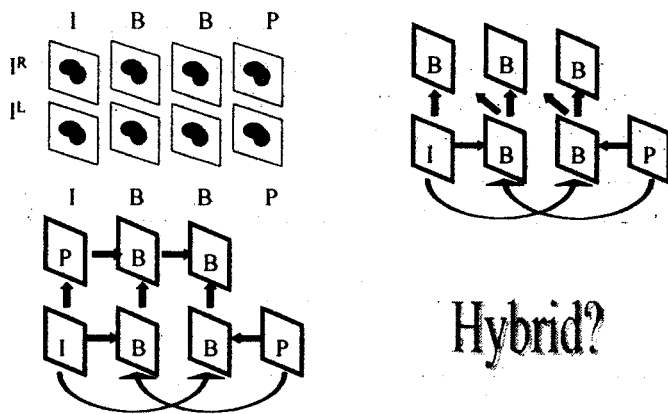
- ◆ BL: codes higher priority bitstream at a lower frame rate
- ◆ EL: codes the intermediate frames

◆ Prediction Configuration



### 3차원 영상압축: Stereo Video Coding

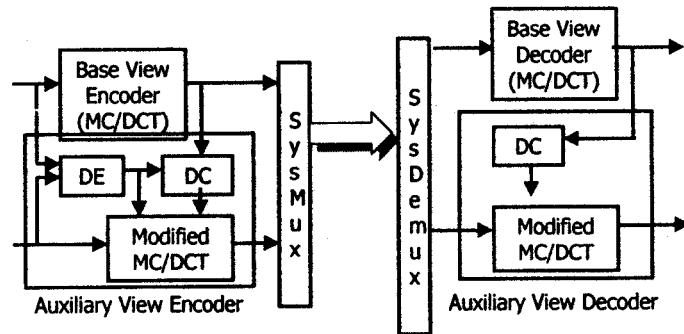
◆ Simulcast vs. Compatible Coding



### 3차원 영상압축: Stereo Video Coding

#### ◆ Structure of Stereo Codec

- ◆ MPEG-2 (13818-3 AMD 3): Multiview Profile (9/9b)
- ◆ Temporal scalability



### 3차원 영상압축: Stereo Video Coding

#### ◆ Experimental Results

- ◆ Input Sequences
  - ◆ CCIR-b01 4:2:2 format with 720x576 at 25Hz (interlaced)
  - ◆ Prediction distance M=3, intraframe distance N=12
  - ◆ Target bit rates: b Mbps + 2 Mbps
- ◆ □ Results (by A. Puri, AT&T)

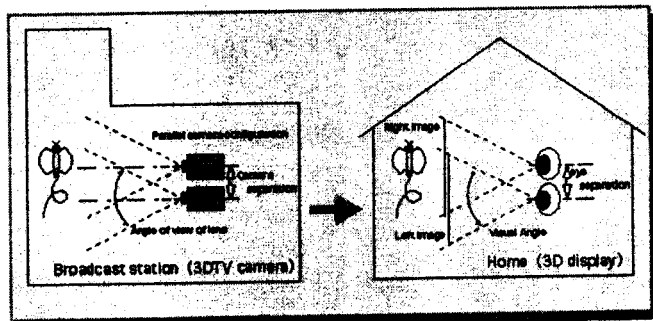
◆ Sequence	Simulcast	DC-DC	DC-MC
◆ Train	31.75	+1.83	+2.81
◆ Manege	26.64	-0.28	+2.20
◆ Tunnel	33.12	-3.17	+1.24
◆ Aqua	30.59	-2.81	+0.50

## Outline

- ◆ Introduction
  - ◆ Motivation and research activities
- ◆ Multiview Image Acquisition
  - ◆ Camera Geometry and 3D Distortion
- ◆ Stereo Video Compression
  - ◆ Scalability tools and Multiview Profile
- ◆ Functionalities for 3D Applications
- ◆ Summary and Conclusion

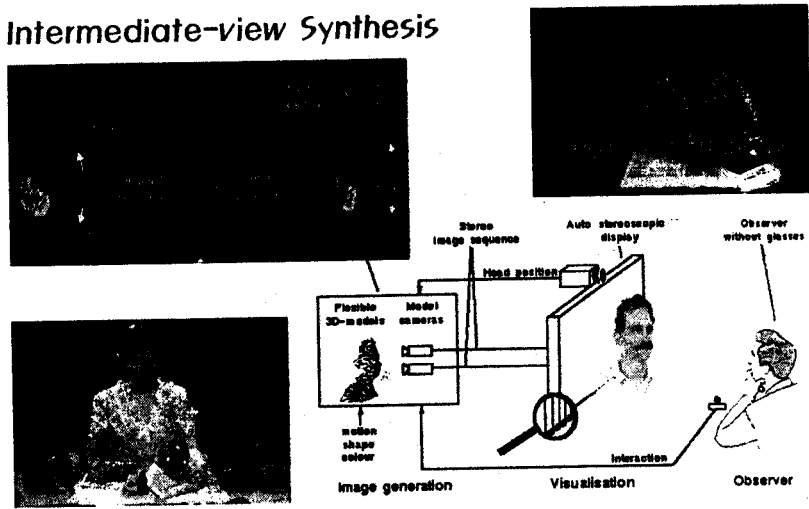
## 3차원 영상처리

- ◆ Why Intermediate-view Synthesis?
  - ◆ 3D HDTV Cameras vs. Human eyes (b<sub>0</sub>-b<sub>5</sub>mm)
  - ◆ Functionality (w/ head tracking)



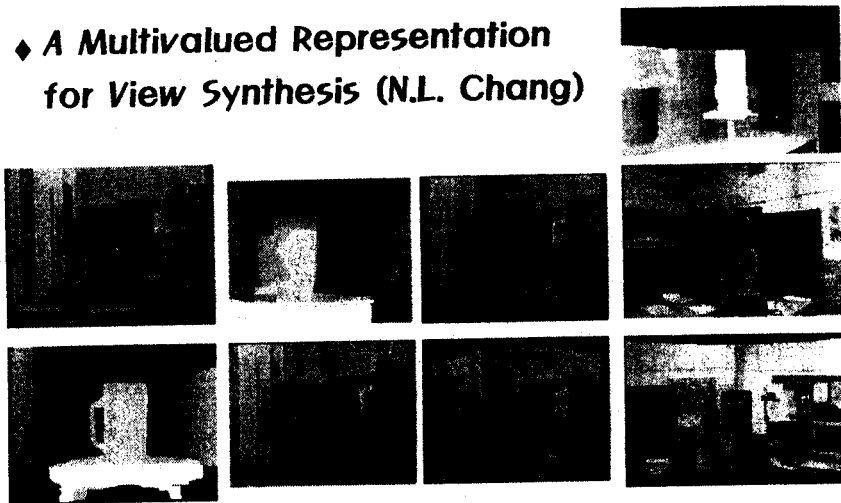
### 3차원 영상처리

#### ◆ Intermediate-view Synthesis



### 3차원 영상처리

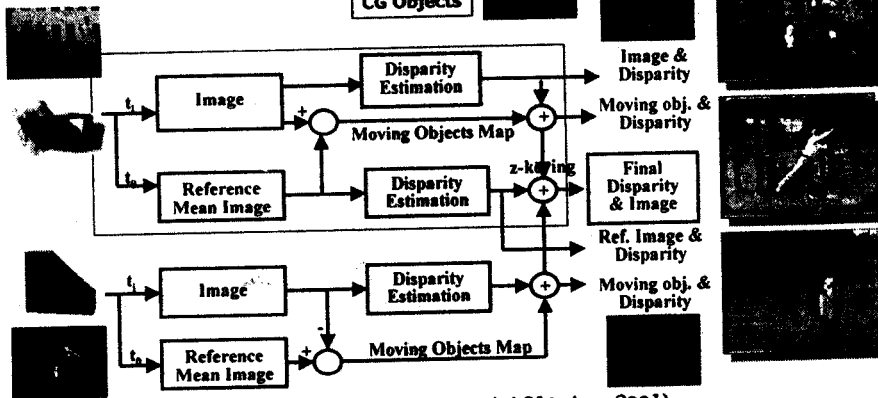
#### ◆ A Multivalued Representation for View Synthesis (N.L. Chang)



### 3차원 영상처리: pi-VE

#### ◆ Photo-realistic interactive VE

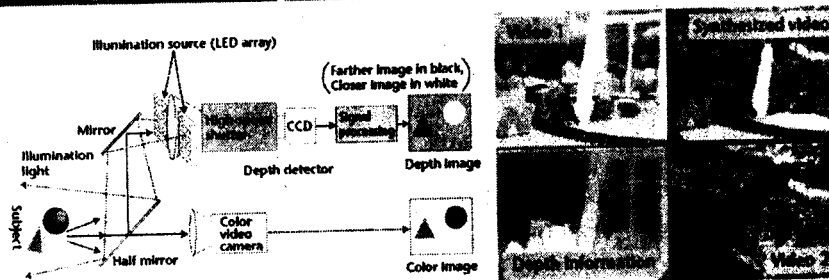
#### ◆ Basic Idea



(in Proc. SPIE VCIP'01, vol. 4310 Jan. 2001)

### 3차원 영상처리

#### ◆ Axi-vision camera (NHK)



## Outline

- ◆ Introduction
  - ◆ Motivation and research activities
- ◆ Multiview Image Acquisition
  - ◆ Camera Geometry and 3D Distortion
- ◆ Stereo Video Compression
  - ◆ Scalability tools and Multiview Profile
- ◆ Functionalities for 3D Applications
- ◆ Summary and Conclusion

## Summary and Q&A

- ◆ Possible Research Topics
  - ◆ 입체영상 영상 촬영 기술 (calibration)
  - ◆ 고화질 입체영상 압축/저장/전송/복원
    - ◆ object-based coding (MPEG-4)
    - ◆ Efficient representation of segmented objects
  - ◆ 대화면 디스플레이용 영상화질 개선 기술
  - ◆ 고화질 입체영상 편집기술
    - ◆ Object-based Functionality:
      - > Background separation
      - > Object-based Segmentation
    - ◆ Flexible Viewing Angle
      - > Multiview video vs. Synthesis of intermediate-view
    - ◆ 입체영상 복원 기술 (VR/MR/AR/Virtual studio)
  - ◆ PUI: personalized 실감 통신
    - ◆ 감성정보처리 기술 & Interactive 영상 정보처리