

# 디지털 비디오 영상 처리 기술

2001. 11.

전병태

(chunbt@etri.re.kr)



한국전자통신연구원  
컴퓨터 \ 소프트웨어연구소

**ETRI**

목 차

- 제1장 디지털 비디오의 특성
- 제2장 디지털 비디오 영상 처리의 필요성
- 제3장 디지털 비디오 영상 처리 기술
- 제4장 원영상 복원 시스템의 개요
- 제5장 복원을 위한 비디오 영상처리 기술
- 제6장 원영상 복원 및 결과

01. 11. 16

무단 복제 및 전제 금지: ETRI ©copyright 2001 ETRI



2

활용 범위의 광 범위성

With advances in computer technology, digital video is becoming more and more common – education, training, entertainment, and publishing.

Data의 용량의 증가성

According to an international survey[1]

- Exist 6 million hours of films
- Video archived worldwide, with a yearly increase rate of about 10% be equal to 1.8 million GB of MPEG
- NASA's Earth Observation System has the capability of generating about 1 terabyte of image data per day

[1] : D.E. Gibson, Report on an International Survey of 500 Audio, Motion Picture Films and Video Archives. Talk given in the annual FIAT/ASA Conference, Sept., 1994. Bogensee, Germany

디지털 비디오 Data 특성

The unique characteristics of video data are discussed by Hampapur[2] and summarized in Table 1-1

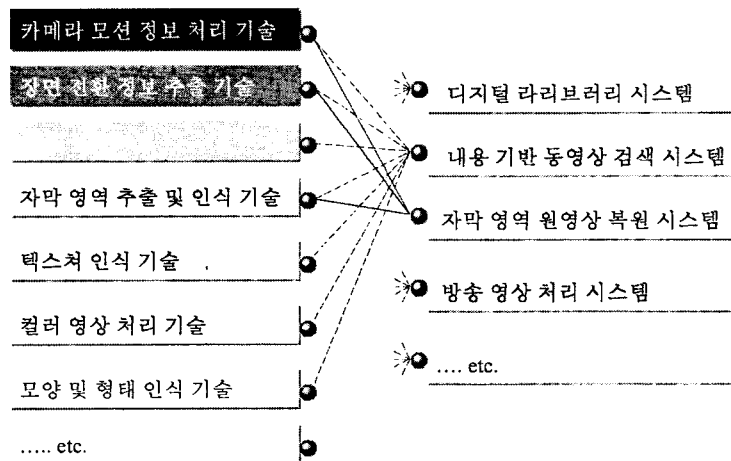
Criteria	Textual Data	Image Data	Video Data
Information	Poor	Rich	Very rich
Dimension	Static and non-spatial	Static and spatial	Temporal and spatial
Organization	Organized	Unstructured	Unstructured
Volume	Low	Median	Massive
Relationship	Simple and well defined	Complex and ill defined	

[2] : Arun Hampapur, Design Video Data Management Systems. PhD thesis The University of Michigan, 1995

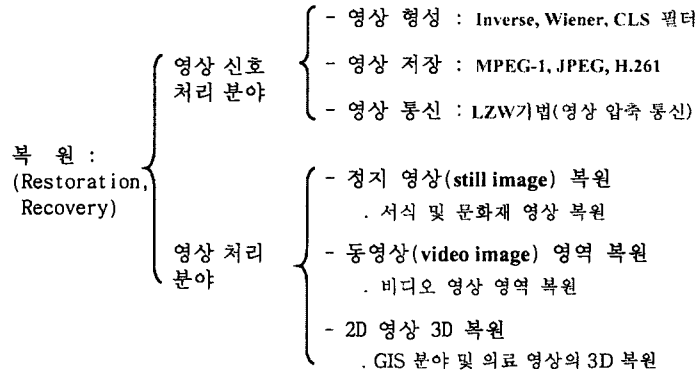
## ETRI 2. 디지털 비디오 영상처리의 필요성

- ▼ 대용량 디지털 비디오 처리 및 활용에 대한 많은 요구 증가  
(예, 내용기반 동영상 검색 시스템)
- ▼ 대용량 영상을 수작업에 의한 처리는 불가능, 자동화된 기술이 필요
- ▼ 영상의 재 사용 및 재활용 기술이 필요

## ETRI 3. 디지털 비디오 영상처리 기술

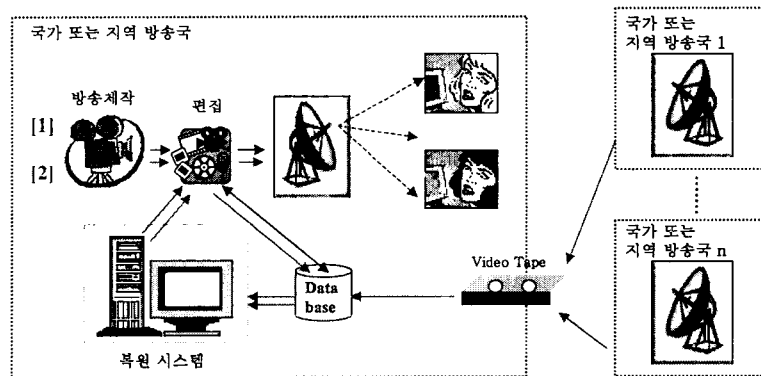


4.1 복원 연구 분류



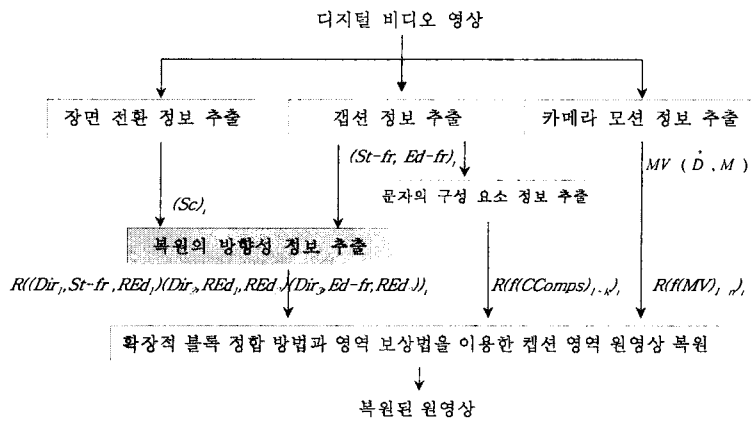
(그림 4.1) 복원의 용어 분류 및 정의

4.2 원영상 복원 시스템



(그림 4.2) 방송 시스템과 연계된 원영상 복원 시스템

## ETRI 4. 원영상 복원 시스템의 개요 (계속)



(그림 4.3) 복원 시스템의 구성도

01.11.16

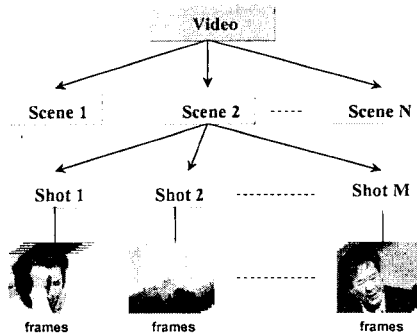
무선 복제 및 전제 금지 ETRI ©copyright 2001 ETRI

9

## ETRI 5. 복원을 위한 비디오 영상 처리 기술

### 5.1 장면 전환 정보 추출 기술

#### 5.1.1 장면 전환의 분류



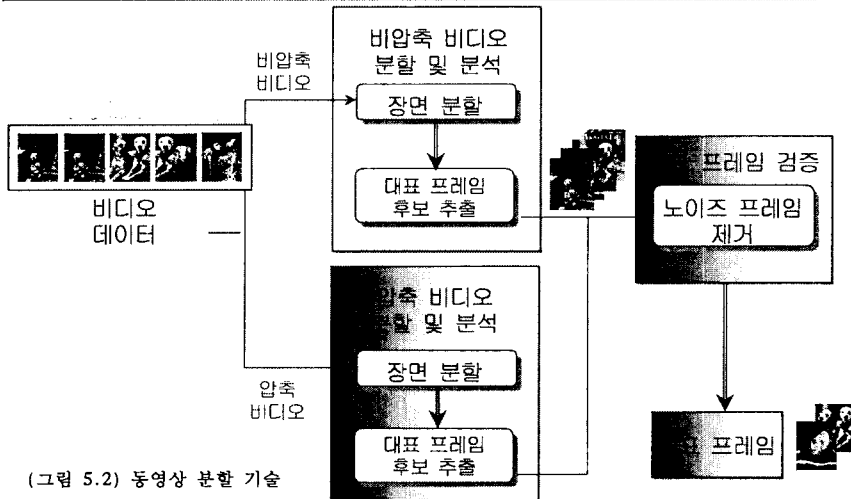
(그림 5.1) 비디오의 구조화

01.11.16

무선 복제 및 전제 금지 ETRI ©copyright 2001 ETRI

10

## ETRI 5. 복원을 위한 비디오 영상 처리 기술 (계속)

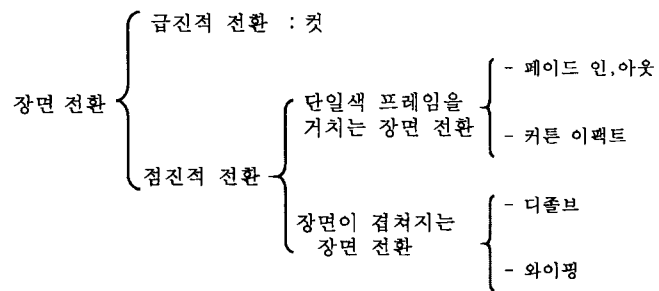


01-11-16

무단 복제 및 전제 금지 ETRI ©copyright 2001 ETRI

11

## ETRI 5. 복원을 위한 비디오 영상 처리 기술 (계속)



(그림 5.3) 급진적 장면 전환 (컷)

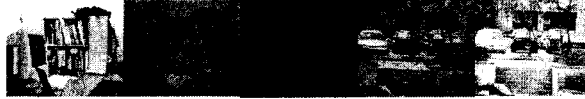
01-11-16

무단 복제 및 전제 금지 ETRI ©copyright 2001 ETRI

12

## ETRI 5. 복원을 위한 비디오 영상 처리 기술 (계속)

(a) 페이딩



(b) 커튼 효과



(c) 디졸브



(d) 와이핑



(그림 5.4) 점진적 장면 전환

## ETRI 5. 복원을 위한 비디오 영상 처리 기술 (계속)

### 5.1.2 장면 전환 검출을 위한 특징량(features)

#### ▼ 화소간 명도차(DOP: Difference of Pixel value)

$$DOP(f_n, f_{n+1}) = \frac{\sum_{i=0}^{N-1} \phi_{f_n, f_{n+1}}(i)}{N}$$

( $f_n, f_{n+1}$ 은 연속된 두 프레임,  $N$ 은 프레임의 크기(가로x세로).  $\phi_{f_n, f_{n+1}}$ 은  $n$ 번째 프레임과  $n+1$ 번째 프레임의 현재 화소의 명도 차를 나타내는 함수로써, 명도차가 임계값 이상이면 1, 그렇지 않으면 0을 갖는다)

#### ▼ 명도 히스토그램 차(DOH: Difference of Histogram)

$$DOH(f_n, f_{n+1}) = \sum_{i=0}^{255} |H_{f_n}(i) - H_{f_{n+1}}(i)|$$

( $H_{f_n}, H_{f_{n+1}}$ 는 각 프레임의 현재 명도 히스토그램 값을 나타내며,  $255$ 는 명도 히스토그램 단계의 개수를 나타낸다)

## ETRI 5. 복원을 위한 비디오 영상 처리 기술 (계속)

### ▼ 차분 영상 히스토그램 분산(HVOD: Histogram variance of difference)

$$HVOD(f_n, f_{n+1}) = \frac{\sum (HOD(f_n, f_{n+1}) - H)}{N}$$

$$HOD(f_n, f_{n+1}) = \frac{\sum_{i \in [-\alpha, \alpha]} hod(i)}{\sum_{i=-Q_0}^{-1} hod(i)} = \frac{\sum_{i \in [-\alpha, \alpha]} hod(i)}{N} \quad hod(i) = h(f_n, f_{n+1})(i)$$

( $HOD_{n,n+1}$ 은  $n$ 번째 프레임과  $(n+1)$ 번째 프레임 사이의 차분영상의 히스토그램,  $N$ 은 차이값이 가지는 범위,  $H$ 은  $H_{n,n+1}$ 의 평균값,  $\alpha$ 는 화소의 변화 여부를 결정하는 임계값으로서, 연속된 두 프레임에서 상응하는 화소 사이의 명도 차의 크기가  $\alpha$  이상인 화소만을 변화가 발생한 것으로 처리하게 된다.  $hod(i)$ 는 연속된 두 프레임의 차영상에 대한  $i$ 번째 명도 히스토그램 값)

### ▼ 화소값의 분산(VOD: Variance of Frame)

$$VOP(f_n, f_{n+1}) = \frac{\sum_{i=1}^K H(i)(i-M)^2}{N}$$

(여기서  $H$ 는 히스토그램,  $K$ 는 화소값의 범위,  $M$ 은 화소수. 한편  $M$ 은 프레임 내의 화소값의 평균인데 이 값도 역시 히스토그램  $H$ 에서 쉽게 구할 수 있다)

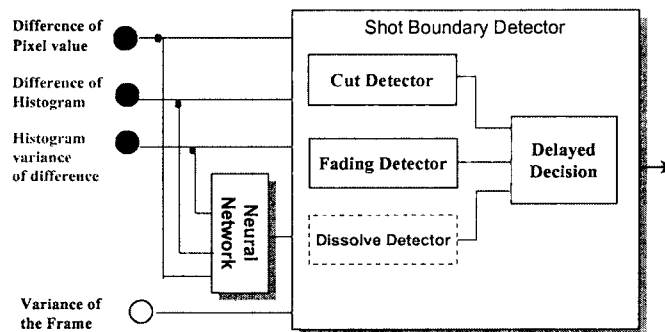
01-11-16

무단 복제 및 전제 금지 ETRI ©copyright 2001 ETRI

15

## ETRI 5. 복원을 위한 비디오 영상 처리 기술 (계속)

### 5.1.3 신경망을 이용한 장면 전환 검출 방법



(그림 5.5) 신경망을 이용한 장면 전환 검출

01 11 16

무단 복제 및 전제 금지 ETRI ©copyright 2001 ETRI

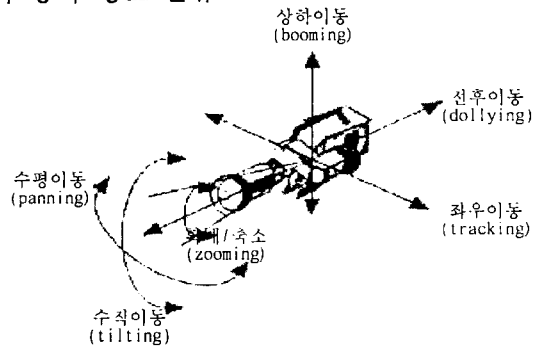
16



## ETRI 5. 복원을 위한 비디오 영상 처리 기술 (계속)

### 5.2 카메라 모션 정보 추출 기술

#### 5.2.1 카메라 동작 정보 분류



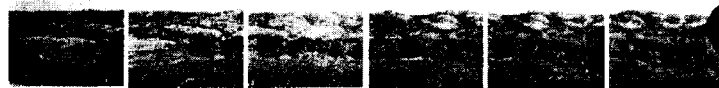
(그림 5.6) 카메라의 기본 동작

01-11-16

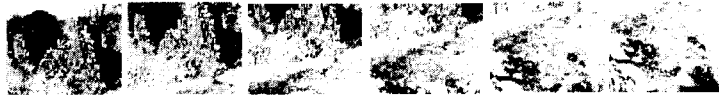
무단 복제 및 전제 금지 ETRI © copyright 2001 ETRI

17

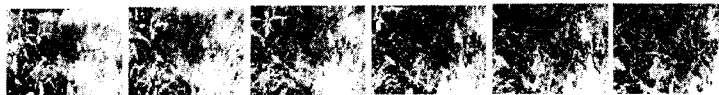
## ETRI 5. 복원을 위한 비디오 영상 처리 기술 (계속)



(a) 수평 이동



(b) 수직 이동



(c) 화면 확대

(그림 5.7) 카메라 동작의 예

01-11-16

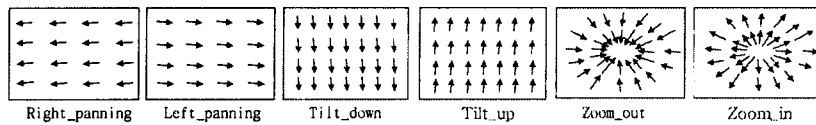
무단 복제 및 전제 금지 ETRI © copyright 2001 ETRI

18

## ETRI 5. 복원을 위한 비디오 영상 처리 기술 (계속)

$$E(dx, dy) = \sum_{y=0}^{N-1-W-1} \sum_{x=0}^{W-1} |f(x, y, t) - f(x-dx, y-dy, t-1)| \dots \text{식(5.1)}$$

( $N$ : 영상의 크기,  $f(x, y, t)$ :  $t$ 시점의  $(x, y)$  좌표의 화소)



(그림 5.8) 카메라 동작 정보 분류

## ETRI 5. 복원을 위한 비디오 영상 처리 기술 (계속)

### 5.2.2 어파인 모델을 이용한 카메라 동작 정보 추출

- ▶ 영상의 움직임은 6개의 파라미터로 구성된 어파인 동작 모델로 표현

$$u(x, y) = a_1 + a_2(x - x_0) + a_3(y - y_0)$$

$$v(x, y) = a_4 + a_5(x - x_0) + a_6(y - y_0) \dots \text{식(5.2)}$$

( $u(x, y), v(x, y)$ 는 영상 좌표  $(x, y)$ 에서 동작 벡터의 수평과 수직 요소)

- ▶ 파라미터  $a_i$  영상의 움직임을 표현

- $a_1$  수평 이동,  $a_4$  수직 이동

- Zoom =  $(a_2 + a_6) = (u_x + v_y)$

- Curl =  $-a_3 + a_5 = -(u_x - v_y)$  ... 식(5.3)

## ETRI 5. 복원을 위한 비디오 영상 처리 기술 (계속)

- 회귀분석법을 이용하여 어파인 동작 모델의 파라미터를 구한다. 예측값과 측정값 사이의 오차를 최소화 할 수 있는 최소제곱법 (method of least squares)을 이용

$$E = \sum_{i=1}^n ((u_i, v_i) - (a_1 + a_2x_i + a_3y_i, a_4 + a_5x_i + a_6y_i))^2$$

$$= \sum_{i=1}^n (u_i - (a_1 + a_2x_i + a_3y_i))^2 + \sum_{i=1}^n (v_i - (a_4 + a_5x_i + a_6y_i))^2 \dots \text{식 (5.4)}$$

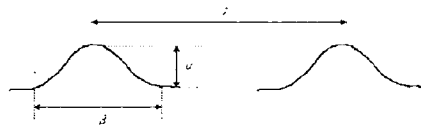
- E를 최소화하는 파라미터  $\mathbf{a} = (a_1, a_2, a_3, a_4, a_5, a_6)$  를 추출위해서는 식(3.4)의 좌변을 파라미터  $a_1, a_2, a_3, a_4, a_5, a_6$  각각에 대해 편 미분 하여 0 으로 놓고 푼다

## ETRI 5. 복원을 위한 비디오 영상 처리 기술 (계속)

### 5.3 자막 영역 추출 기술

#### 5.3.1 자막 영역 추출 방법

- 라인 샘플링 방법과 문자의 지형학적 특징을 이용

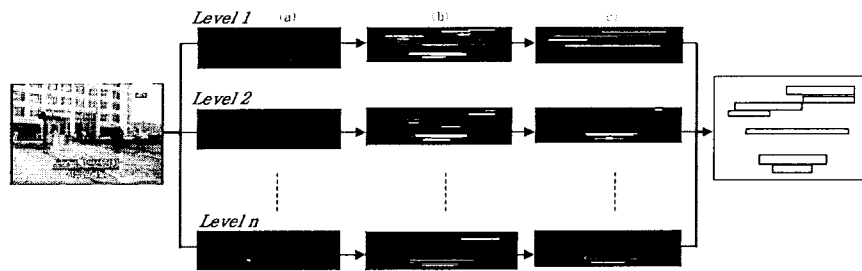


(그림 4.8) 문자의 지형학적 특성 및 상호 관계

$$TF_{ij} = \begin{cases} 1 & : \text{if } ((P_{ij} (= A_1) > A_2 + \alpha_{th}) \&\& ((A_2 > A_3 + \alpha_{th}) \&\& \dots \\ & \dots (A_{n-1} > A_n + \alpha_{th})) \\ 0 & : \text{otherwise} \end{cases} \dots \text{식 (5.5)}$$

## ETRI 5. 복원을 위한 비디오 영상 처리 기술 (계속)

### 5.3.2 문자의 후보 영역 추출 방법

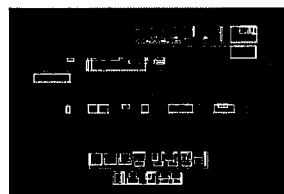


(그림 5.9) 캡션 후보 영역 추출 과정

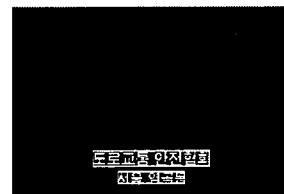
## ETRI 5. 복원을 위한 비디오 영상 처리 기술 (계속)

### 5.3.3 문자의 후보 영역 검증 방법

- ▶ 문자의 후보 영역 이치화
- ▶ 대단위 영역화를 통한 후보 영역 검증



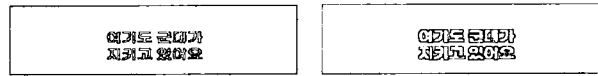
(그림 5.10) 대단위 영역화 방법에 의한 검증



(그림 5.11) 검증 후 추출된 캡션 영역

6.1 디지털 비디오 자막의 특성

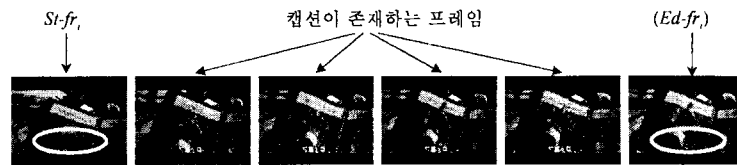
- 복원 될 문자 영역은 실제 캡션 영역보다 크다



(그림 6.1) 캡션 문자의 특성

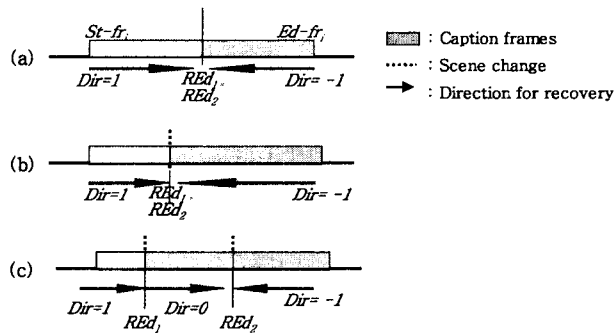
- 비디오 캡션은 대화나 설명을 위한 장면에 주로 삽입  
다른 장면에 비하여 물체의 움직임이 적고 카메라 동작이 비교적 단순함을 볼 수 있다.  
캡션 영역에서의 장면 전환은 자주 발생하지 않는다

- 캡션이 시작되는 전 프레임 (*St-fr*)이나 캡션이 끝나는 다음 프레임 (*Ed-fr*)은 캡션이 삽입되지 않는 원 영상을 갖고 있다



(그림 6.2) 비디오 프레임의 특성

6.2 원영상 복원의 방향성 결정



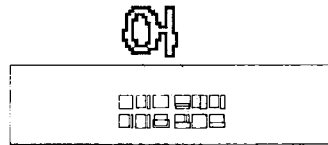
(그림 6.3) 원영상 복원의 방향성 결정

6.3 캡션 영역의 원영상 복원 방법

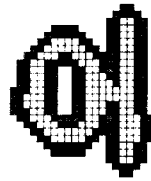
6.3.1 원영상 복원 구조

- ▶ 복원의 방향성 ( $Dir_j$ )에 따라 복원 방법 결정
- ▶ 추출된 문자의 구성요소 단위로 원영상 복원 수행
- ▶ 확장적 블록 정합 방법, 영역 보상법을 이용 원 영상 복원
- ▶ 복원 종료 후, 역 방향에서 복원된 영상의 재 정렬

6.3.2 구성 요소 단위의 복원



(그림 6.4) 문자의 구성 요소

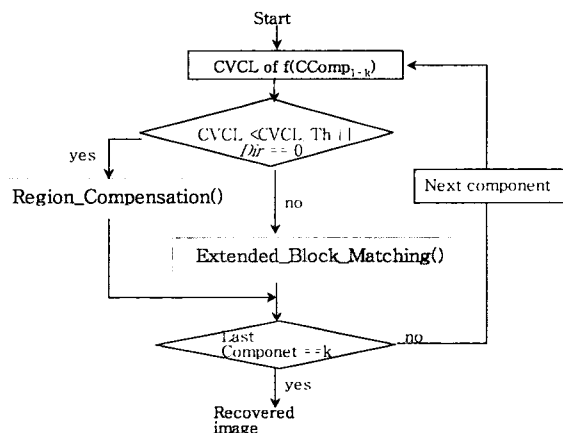


(그림 6.5) 복원될 문자 영역

- : 원 영상 외곽선
- : 문자 그림자 선
- : 문자 영역

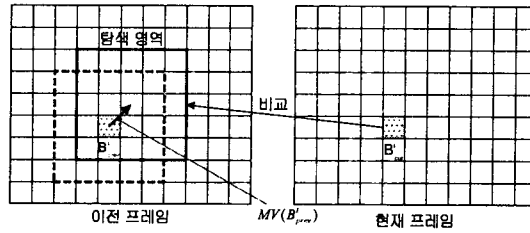
$$CVCL = \frac{\sum_{i=0}^N \left( \sqrt{(Rc - R_i)^2 + (Gc - G_i)^2 + (Bc - B_i)^2} \right)}{N} \quad \dots \text{식(6.1)}$$

$$Rc = \frac{\sum_{i=0}^N R_i}{N} \quad Gc = \frac{\sum_{i=0}^N G_i}{N} \quad Bc = \frac{\sum_{i=0}^N B_i}{N}$$



(그림 6.6) 각 구성 요소 단위의 원영상 복원 과정

6.3.3 확장적 블록 정합(Extended Block Matching) 방법



$MV(B^i_{pre})$  : 초기 분할된  $B^i_{cur}$  블록과 이전 시점의  $B^i_{prev}$  블록의 동작 벡터의 합

(그림 6.7) 탐색 영역의 동적인 위치 설정

- 동작 벡터의 신뢰도  $CF(MV(B^i_{pre}))$ 를 구함

$$CF(MV(B^i_{pre})) = \frac{K_1}{1 + K_2 \cdot VD(B^i_{pre})} \dots \text{식(6.2)}$$

$$VD(B^i_{pre}) = \|MV(B^i_{pre}) - \mu\|^2 \cdot \sigma^2$$

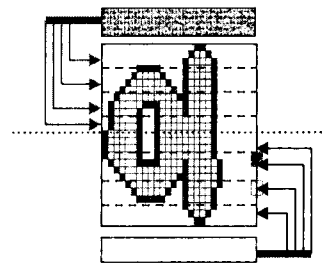
- $VD(B^i_{pre})$  :  $B^i_{pre}$  블록과 인접된 블록들에서 구한 동작 벡터 사이의 분산이 보상된 거리
- $K_1$  : 신뢰도가 취할 수 있는 최대값을 결정
- $K_2$  : 신뢰도의 범위를 결정
- $\mu$  = mean of  $MV(B^i_{pre})$ ,  $i^* \in \text{neighborhood of } i$
- $\sigma^2$  = variance of  $MV(B^i_{pre})$ ,  $i^* \in \text{neighborhood of } i$

- 신뢰도에 따라 탐색 범위를 조정. 신뢰도가 높으면 탐색 범위를 좁히고, 적으면 탐색 범위를 크게 한다



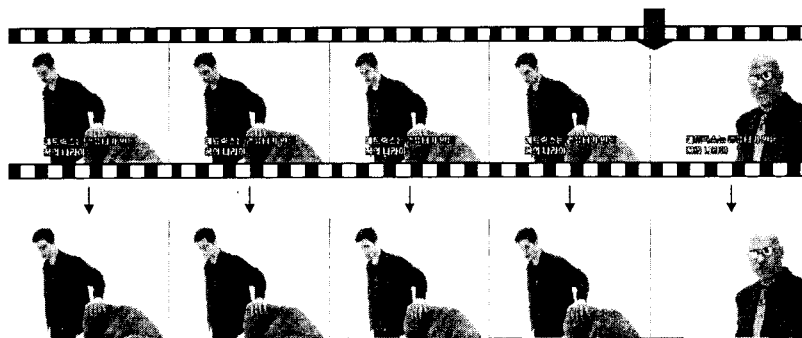
6.3.4 영역 보상법 (Region compensation)

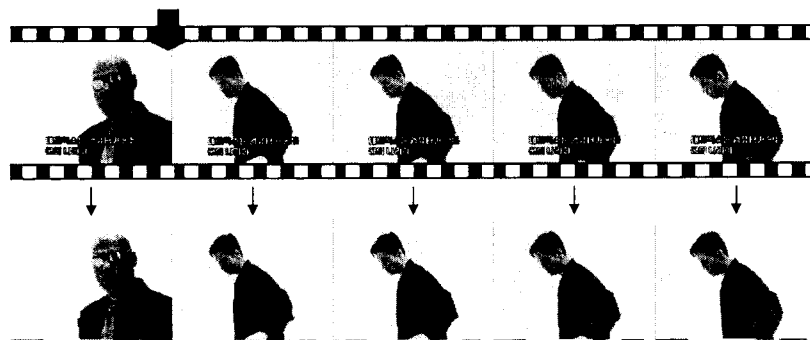
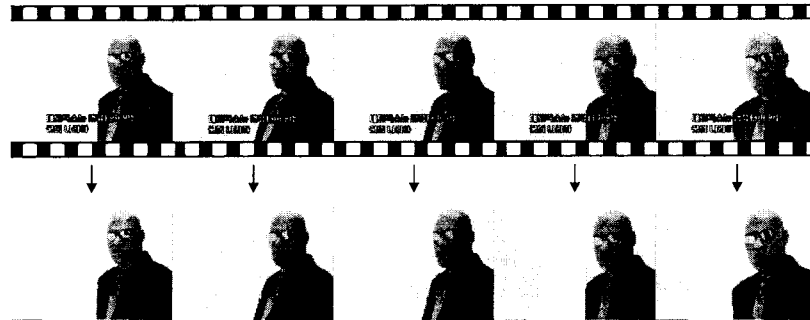
- ▶ 캡션 영역을 상하 2 등분한다
- ▶ 상단에서 보상 영역을 구한다
- ▶ 보상 영역 단위로 캡션 영역에 대한 복원을 반복
- ▶ 하단에 대하여도 동일한 방법 수행



(그림 6.8) 영역 보상법

6.4 영화(“matrix”)에 대한 실험 결과





(그림 6.9) 장면이 2회 이상 발생된 경우



(그림 6.10) 기존 자막 변경 방법과의 비교