

# 확장적 블록 정합 방법과 영역 보상법을 이용한 비디오 문자 영역 복원 방법

## (A Method for Recovering Text Regions in Video using Extended Block Matching and Region Compensation)

전 병 태 <sup>†</sup> 배 영 래 <sup>\*\*</sup>  
(Byung Tae Chun) (Younglae Bae)

**요약** 기존의 원영상 복원 기술은 주로 신호 처리 분야에서 영상의 형성(formation), 저장 및 통신상에서 발생하는 왜곡 현상을 복원하는 연구가 많이 이루어 졌다. 원 영상 복원과 관련된 기존의 연구는 블록 정합(block matching algorithm)을 이용한 원영상 복원 방법이 있다. 이 방법은 오 정합(incorrect matching) 발생하기 쉽고 오 정합 시 에러가 전파되는 문제점이 있다. 그리고 장면 전환이 2회 이상 발생될 경우 장면 전환 지점과 지점 사이의 복원이 불가능하다는 문제점이 있다. 본 논문에서는 기존의 문제점들을 해결하기 위하여 확장적 블록 정합 방법(EBMA: Extended Block Matching Algorithm)과 영역 보상법(Region compensation method)을 이용한 원영상 복원 방법을 제안하고자 한다. 원영상 복원에 사용하기 위하여 비디오 사전 정보(장면 전환 정보, 카메라 모션 정보, 캡션 영역 정보)를 추출한다. 추출된 캡션 영역 정보를 이용하여 캡션 문자의 구성 요소 정보를 추출한다. 추출된 비디오 사전 정보를 이용하여 복원의 방향성을 결정하고, 복원의 방향성에 따라 문자의 구성 요소 단위로 확장적 블록 정합 방법과 영역 보상법을 이용하여 원영상 복원을 수행한다. 실험결과 확장적 블록 정합 방법은 빠른 물체의 움직임이나 복잡한 배경에 영향을 받지않고 복원이 잘 되는 것을 볼 수 있었다. 참조할 원영상이 없이 원영상 복원을 수행하는 영역 보상법의 복원 결과 또한 좋음을 볼 수 있었다.

**키워드** : 자막 추출, 문자영역복원, 비디오 영상, 영역보상법, 확장적 블록정합방법

**Abstract** Conventional research on image restoration has focused on restoring degraded images resulting from image formation, storage and communication, mainly in the signal processing field. Related research on recovering original image information of caption regions includes a method using BMA(block matching algorithm). The method has problems with frequent incorrect matching and propagating the errors by incorrect matching. Moreover, it is impossible to recover the frames between two scene changes when scene changes occur more than twice. In this paper, we propose a method for recovering original images using EBMA(Extended Block Matching Algorithm) and a region compensation method. To use it in original image recovery, the method extracts a priori knowledge such as information about scene changes, camera motion and caption regions. The method decides the direction of recovery using the extracted caption information(the start and end frames of a caption) and scene change information. According to the direction of recovery, the recovery is performed in units of character components using EBMA and the region compensation method. Experimental results show that EBMA results in good recovery regardless of the speed of moving object and complexity of background in video. The region compensation method recovered original images successfully, when there is no information about the original image to refer to.

**Key words** : text extraction, recovering text region, video images

<sup>†</sup> 정 회 원 : 한국전자통신연구원 컴퓨터·소프트웨어 연구소  
chunbt@etri.re.kr

<sup>\*\*</sup> 비 회 원 : 충북과학대학 전자상거래학과 교수  
yjb@ctech.ac.kr  
논문접수 : 2002년 12월 17일  
심사완료 : 2002년 8월 22일

## 1. 서론

멀티미디어 기술은 다 매체를 효율적으로 처리하여 사용자에게 이용하기 편리한 최적의 형태를 제공하는 기술이라고 볼 수 있다. 멀티미디어 기술에 사용되는 다 매체

중 동영상 매체(예, 디지털 비디오)의 사용은 날로 증가하는 추세이며, 현재 동영상 처리 및 관리에 관한 연구가 활발히 진행되고 있다. 최근의 연구는 디지털화 된 수많은 영상 및 비디오 정보를 활용 및 재 사용하는 측면을 많이 생각하게 되었다. 저장 자료의 사용성을 높이기 위해서 저장된 비디오 영상을 빠른 시간에 검색할 수 있는 동영상 검색 시스템이나 원 영상 복원 시스템과 같은 기술이 개발되고 있다.

영상 복원 분야를 분야별로 분리하여 보면 그림1과 같이 영상 신호 처리 분야와 영상 처리 분야로 나누어 볼 수 있다. 대부분 영상 복원에 관한 연구는 영상 신호 처리 분야에서 많이 이루어 졌다. 영상 신호 처리 분야에서 영상의 형성, 저장 및 전송과 관련된 일련의 과정에서 변형 또는 왜곡에 의하여 변화된 영상을 복원하는 연구이다.

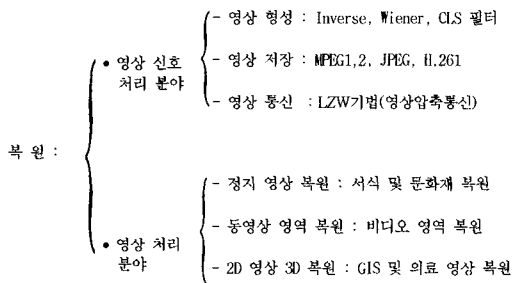


그림 1 영상 복원에 대한 기존 연구의 분류

영상 처리 분야에서의 복원은 형성된 영상에서 특정 영역 또는 훼손된 영역에 대한 복원 또는 재구성을 의미한다.

영상 처리 분야에서 동영상 복원의 필요성, 특히 비디오 자막 영역의 원영상 복원의 필요성은 다음과 같은 3가지 이유에서 필요하다고 볼 수 있다. 첫째, 복원하고자 하는 영상의 양이 적을 경우 수 작업 처리가 가능하나 동영상의 경우처럼 처리하여 할 영상의 양이 많을 경우 수 작업에 의한 복원은 처리 시간 많이 소요됨으로써 복원이 불가능 할 수 있다. 따라서 자동화된 복원 기술이 필요하다. 둘째, 영상의 재사용이라는 관점에서 반드시 필요한 기술이라고 볼 수 있다. 방송국에서 자막이 삽입된 영상을 필요에 따라 원 영상으로 복원하여 재 사용할 필요성이 발생할 경우, 자막 영역을 원 영상으로 복원하여 방송함으로써 시청자에게 시청의 편의성을 제공 할 수 있다. 셋째, 외국에서 제작된 영상에 외국어 자막이 삽입되었을 경우 방송물에 삽입된 자막을 제거하고 원 영상으로 복원한 후 원하는 언어 자막으로

대치함으로써 다른 자막으로 대치의 편리성을 제공할 수 있다.

본 논문에서는 원영상 복원에 필요한 비디오 사전 정보(prior-knowledge) 추출하고 확장적 블록 정합 방법과 영역 보상 기법을 이용하여 자막 영역에 대한 원영상 복원 방법을 제안하고자 한다. 추출된 비디오 사전 정보(장면 전환 정보, 자막 추출 정보)를 이용하여 복원의 방향성을 결정하고, 복원의 방향성 정보와 카메라 동작 정보를 이용하여 자막의 구성 요소 단위로 복원을 진행한다. 전체적인 시스템의 구조는 그림2와 같다.

논문의 구성은 6장으로 구성된다. 1 장에서는 복원의 용어 정의 및 필요성에 대하여 살펴 본다. 2장에서는 기존의 연구 방법에 관하여 살펴보고, 3장에서는 복원에 필요한 비디오 사전 정보 추출 방법에 관하여 간략히 살펴본다. 그리고 삽입된 자막의 특성에 대하여 설명한다. 4장에서는 영역 보상법과 확장적 블록 정합방법을 이용 비디오 자막 영역을 원영상 복원 방법에 대하여 설명한다. 5장에서는 영화“매트릭스”, “마야”에 대하여 원영상 복원 실험 결과를 보여주고, 6장 결론에서는 향후 연구 방향에 대하여 설명한다.

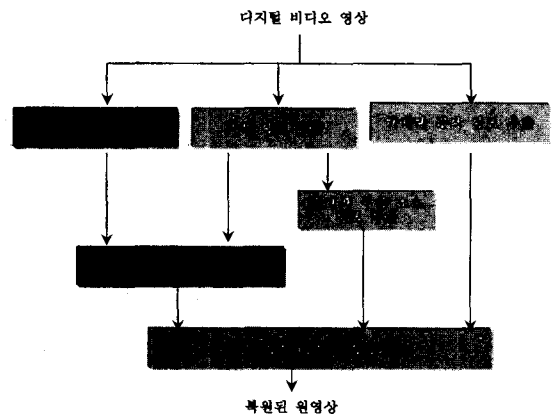


그림 2 비디오 자막 영역의 원영상 복원 방법

## 2. 기존의 연구 방법 고찰

정지 영상에서 사용되는 국보급 문화재 복원 방법[1]으로는 주로 보간법(interpolation)을 이용한다고 볼 수 있다. 보간법에 의한 원영상 복원은 선 위주의 복원으로 캡션 영역처럼 복원 영역이 큰 복원에는 적합하지 않을 뿐 아니라 또한 복원된 결과가 질감(texture)을 표현하지 못 함으로서 블록화 현상이 발생하는 문제점이 있다.

디지털 비디오 원영상 복원에 관한 연구는 블록 정합

방법을 이용한 원영상 복원 방법이[2, 3, 4, 5]있다. 이 방법은 블록 정합 방법을 이용하여 복원하려는 문자 영역과 원영상과의 정합을 수행, 최소값을 갖는 영역을 찾아 원 영상으로 복원 하는 방법이다. 이 방법의 문제점을 3가지로 볼 수 있다. 첫째, 단순한 블록 정합 방법을 이용함으로써 물체와 카메라의 급격한 이동으로 발생하는 문제를 수용할 수 없고 그로 인해 오 정합(incorrect matching)이 발생할 수 있다. 둘째, 오 정합 발생시 다음 프레임에 영향을 줌으로써 오류가 계속 전파(propagation)될 수 있다는 문제점이 있다. 셋째, 장면 전환이 2회 이상 발생할 경우, 장면 전환 지점과 장면 전환 지점사이에서는 참조할 원 영상이 없으므로 이 부분에 대해서는 처리할 수 없다는 문제점이 있다. 본 논문에서는 기존 연구 방법의 문제점을 보완하기 위하여 확장적 블록 정합 방법과 영역 보상법을 이용하여 기존의 문제를 보완하려고 한다.

### 3. 디지털 비디오 사전 정보 추출 및 삽입된 자막의 특성

#### 3.1 장면전환 정보 추출

동영상 장면 전환 정보는 어디서 장면 전환이 발생되었는지를 찾아 이 정보를 원영상 복원에 이용한다. 여기서 장면이란 비디오 데이터 분할의 기본 단위로 끊기지 않고 촬영된 일련의 영상들로서 시공간적으로 연속적인 행동을 보여주는 영상 시퀀스를 의미한다. 장면 전환이 발생하는 지점의 전 · 후 프레임은 서로 관련성이 없는 다른 영상으로 볼 수 있으므로 장면 전환 지점은 처리 단위를 나누기 위한 기준점이 될 수 있으며 복원을 위한 방향성 정보 결정시 중요한 정보가 될 수 있다. 기존의 연구 방법[6,7,8]은 위의 두 문제를 동시에 해결하기 위한 방법을 모색하려는 연구가 이루어졌다.

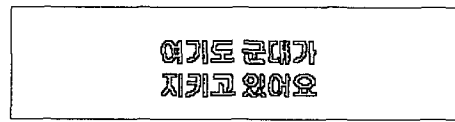
#### 3.2 카메라 모션 정보 추출

카메라 동작 정보는 카메라 동작의 종류와 크기에 관한 정보를 추출하는 것이다[9, 10, 11]. 즉, 카메라 동작 정보를 추출함으로써 물체 또는 배경 영역에 대한 움직임의 종류 및 크기를 알 수 있고, 이 정보는 확장적 블록 정합 시 복원의 참조 위치 결정에 중요한 정보가 된다.

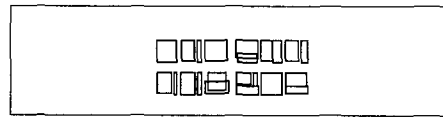
#### 3.3 비디오 자막 영역 추출 및 문자의 구성요소 정보 추출

자막 정보는 복원하고자 하는 자막 영역이 어디에 위치하는지에 대한 정보(프레임 번호, 문자 영역 좌표)를 추출하는 것이다[12,13]. 자막 정보 추출은 자막 영역 원영상 복원에 중요한 정보라고 볼 수 있다. 자막의 위

치와 색상은 대부분 고정적이지만 항상 일정하다고 볼 수 없으므로 위치와 색상에 의존하지 않는 자막 영역 추출 방법[14]을 사용하여야 된다. 추출된 자막의 프레임 번호 정보는 복원의 방향성 정보 결정에 사용되며, 추출된 문자 영역에서 문자의 구성 요소 정보를 추출에 사용한다. 문자의 구성 요소 정보란 문자의 자소 단위를 말하며 원영상 복원의 기본 단위로 사용된다. 그림 3(a)는 추출된 자막 영역을 보여주고 있으며, 그림3(b)는 그림 3(a)에서 구해진 자막의 구성요소 정보를 보여 주고있다.



(a) 추출된 자막 영역



(b) 문자의 구성 요소 추출

그림 3 복원의 기본 단위(문자의 구성요소 정보)

#### 3.4 비디오 자막 영역의 특성

비디오에 삽입된 자막 영역은 다음과 같은 몇 가지 특성을 갖고 있으며, 본 논문에서는 그 특성을 이용하여 자막 영역을 원 영상으로 복원에 한다. 비디오에 삽입된 자막 문자의 특성은 다음과 같이 요약할 수 있다.

첫째, 복원할 자막 영역은 실제 자막 문자 영역보다 크다. 비디오에 삽입된 자막 문자는 문자 영역을 강조하기 위하여 그림 4와 같이 문자 주위에 그림자 선을 삽입한다. 이러한 문자의 윤곽선 영역 역시 원 영상을 손상시켰다고 볼 수 있으므로 원 영상 복원 시 복원에 참여 시켜야 한다. 그렇지 않을 경우 문자의 그림자 선 부분이 복원되지 않는 문제가 발생한다.

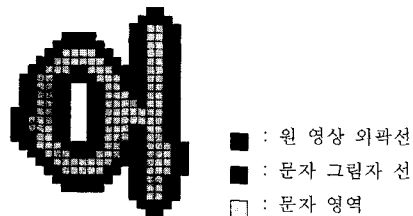


그림 4 문자의 구성요소 영역과 관련된 영역

둘째, 비디오 자막은 주로 대화나 설명을 위한 장면에 주로 삽입된다. 따라서 다른 장면에 비하여 물체(사람 또는 사물)의 움직임 비교적 적고 카메라의 움직임이 비교적 단순함을 볼 수 있다.

셋째, 자막이 시작되는 전 프레임이나 자막이 종료는 다음 프레임은 그림 5와 같이 자막이 삽입되지 않는 원 영상을 갖고 있다. 이러한 원 영상은 복원을 위한 중요한 시작 정보로 사용한다.

넷째, 자막이 삽입된 부분에서는 장면 전환이 자주 발생하지 않는 특성이 있다. 왜냐하면, 3번째 가정에서 언급했듯이 대부분의 자막 영역은 대화 및 설명 장면이므로 많은 장면 전환은 자주 발생하지 않음을 볼 수 있다. 장면 전환이 발생한다면 대부분 1 ~ 2회 정도 발생하며, 3 번 이상의 장면 전환은 거의 발생하지 않는다.

다섯째, 자막은 비디오의 내용을 시청자에게 보다 더 정확한 의미 전달의 목적으로 삽입되며, 대부분 영상의 하단에 삽입된다. 그리고 자막의 생성과 소멸 시간은 최소 1 ~ 7초 범위에서 자막이 생성되었다 소멸된다. 따라서 자막 영역 추출의 최소 프레임 비율 1 frame/1 sec 정할 수 있다.

#### 4. 확장적 블록 정합 방법을 이용한 원 영상 복원 방법

##### 4.1 원영상 복원의 방향성 결정

복원의 방향성을 결정하는 이유는 그림 5와 같이 자막이 시작되는 전 프레임이나 자막이 끝나는 다음 프레임은 자막이 삽입되지 않는 원 영상을 갖고있기 때문이 이 영상을 원 영상 복원에 중요한 정보로 이용하려는 것이다. 추출된 자막 정보(시점과 종점 정보) 및 장면 분할 정보는 원 영상 복원의 출발점과 복원의 방향성을 결정하는 중요한 정보로 이용된다.

추출된 자막 정보와 장면 전환 정보를 이용하여 복원의 방향성을 결정하며, 어느 방향으로 시작하여 어디까지

복원을 진행할 것인가를 결정한다. 복원의 방향성은 3가지로 분류 할 수 있다.

첫째, 자막 영역에 장면 분할이 발생치 않을 경우, 자막 복원의 방향성은 자막의 시작 전 프레임부터 중앙 지점까지 순방향으로 복원하고, 자막 종료 다음 프레임에서 중앙 지점까지 역 방향으로 복원을 한다. 여기서 중앙 지점은 자막의 시작과 종료의 구간의 중앙 지점을 말한다.

둘째, 장면 분할이 1 회 발생 할 경우, 시작 전 프레임에서 시작하여 장면 분할 지점까지 순 방향으로 복원 하고, 자막 종료 다음 프레임에서 시작하여 장면 분할 지점까지 역 방향으로 복원한다.

셋째, 장면 분할이 2 회 이상 발생하는 경우, 자막이 존재하는 시작 프레임 전 프레임부터 시작하여 첫 장면 분할 지점까지 순방향으로 복원 하며, 자막 종료 다음 프레임부터 시작하여 두 번째 장면 전환 지점까지 역 방향으로 복원한다. 장면 분할 지점과 지점 사이는 순방향으로 복원을 진행한다.

#### 4.2 확장적 블록 정합 방법과 영역 보상법을 이용한 원영상 복원 방법

##### 4.2.1 복원될 문자의 특성에 기반한 복원 방법 결정

그림 4의 문자의 외곽선 화소에 대한 컬러 분산(CVCL: Color Variance for Contour line) 값을 구한다. 구해진 CVCL 값이 임계값(CVCL\_Th) 이상이라고 할 경우, 복잡한 배경 영역이 걸쳐 있거나 경계선이 단일 색상이 아닐 경우로 볼 수 있다. 만약 CVCL 값이 임계값 이하라고 하면 문자 주변의 배경 영역 색상이 일정하다고 볼 수 있다. 즉, 단일 색상일 가능성이 많다. 따라서 본 논문에서는 CVCL 값을 구하여 복원될 문자 영역의 성격을 파악하고 적절한 복원 방법을 선택하여 원영상으로 복원하는 방법을 선택한다. CVCL 값을 구하는 식은 식(1)와 같다.

$$CVCL = \frac{\sum_{i=0}^N (\sqrt{(Rc - Ri)^2 + (Gc - Gi)^2 + (Bc - Bi)^2})}{N} \quad (1)$$

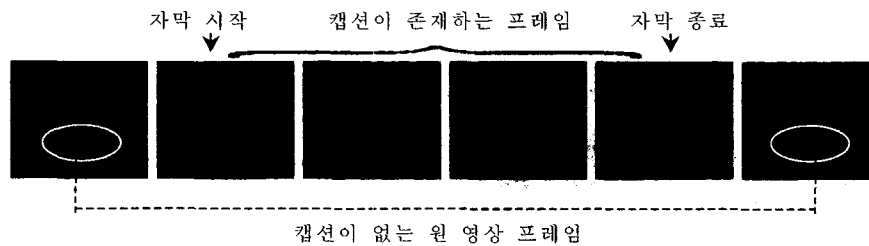


그림 5 자막의 시작 및 끝 프레임

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

CVCL값과 장면 전환 정보에 따라 복원 방법을 그림 6처럼 달리 할 수 있다.

구해진 CVCL 값이 임계값 이하라면 단일 색상으로 볼 수 있으므로 영역 보상법을 사용하여 문자의 구성요소 영역을 복원한다. CVCL값이 임계값 이상이고 장면 전환이 1회 이하 발생하였으면 확장적 블록 정합 방법을 이용하여 복원하고, 장면 전환이 2회 이상 발생하였을 경우 문자 영역 보상법을 이용하여 복원한다.

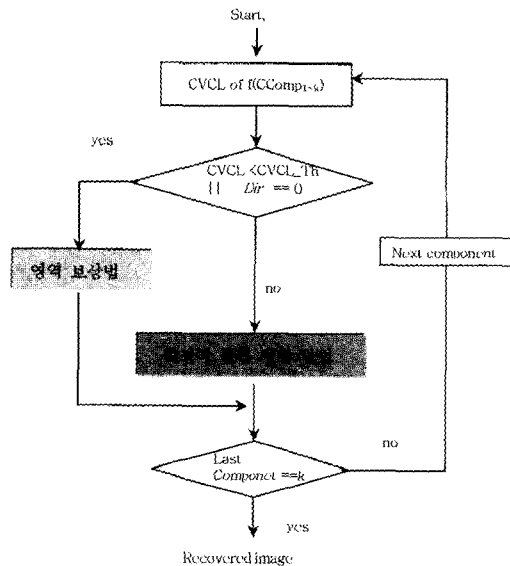


그림 6 복원될 문자의 구성 요소 단위의 복원 방법

#### 4.2.2 확장적 블록 정합 방법

블록 정합은 영상을 모양과 크기가 동일한 사각형의 블록으로 분할한 후 블록 단위의 특징을 정의하고, 정의된 특징에 의해 구성되는 정합 척도를 일정한 탐색 영역 내의 블록들에 적용하여 가장 높은 정합 척도를 가지는 블록을 찾는다[15]. 확장적 블록 정합 방법의 탐색 영역 크기는 식(2)와 같이 정의된다[16][17]. 탐색 영역의 크기는 동작 벡터의 신뢰도에 따라 연속적으로 변화하며, 동작 벡터의 신뢰도는 탐색 영역의 크기를 확대 또는 축소시키는 가중치의 역할을 하도록 정의된다. 다시 말해서, 이전 시점에서 추출된 동작 벡터가 높은 신뢰도를 가질 경우에는 블록 정합이 정확하고 안정적으로 수행되었다는 것을 나타내므로 현재 시점에서 탐색 영

역의 크기를 축소하여 시간적인 복잡도를 최소화하고, 이전 시점에서 추출된 동작 벡터가 낮은 신뢰도를 가질 경우에는 블록 정합이 불안정하게 수행되었다는 것을 나타내므로 현재 시점에서 탐색 영역의 크기를 확대하여 블록 정합의 정확도를 최대화할 수 있도록 한다.

$$s_{B_{cur}}(x, y) = s_{min}(x, y) + (1 - CF(MV(B_{pre}^i))) \times (s_{max}(x, y) - s_{min}(x, y)) \quad (2)$$

식(2)에서 탐색 영역의 크기  $s_{B_{cur}}(x, y)$ 는 탐색 영역의 최소 크기와 최대 크기를 나타내는  $s_{max}(x, y)$ 과  $s_{min}(x, y)$  및 블록  $B_{pre}^i$ 에서 추출되는 동작 벡터의 신뢰도  $CF(MV(B_{pre}^i))$ 를 이용하여 정의된다.

$B_{cur}^i$ 는 현재 시점의  $i$  번째 블록,  $MV(B_{pre}^i)$ 는 이전 시점에서 추출한 블록  $B_{pre}^i$ 의 동작 벡터를 나타낸다.

이러한 개념을 기반으로 동작 벡터의 신뢰도를 식(3)와 같이 정의한다. 이전 시점의 동작 벡터의 신뢰도는 블록  $B_{pre}^i$ 에서 구한 동작 벡터와 블록  $B_{pre}^j$ 의 인접 블록들에서 구한 동작 벡터들을 비교하여 구한다. 일반적으로, 인접한 블록들은 유사한 동작 벡터를 가질 확률이 높으므로 블록  $B_{pre}^i$ 에서 구한 동작 벡터와 인접 블록들에서 구한 동작 벡터들의 차이가 클수록 블록은 큰 움직임을 가질 확률이 높고, 동작 벡터들의 차이가 작을수록 블록은 작은 움직임을 가질 확률이 높다.

$$CF(MV(B_{pre}^i)) = \frac{K_1}{1 + K_2 \cdot VD(B_{pre}^i)} \quad (3)$$

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

$\mu = \text{mean of } MV(B_{pre}^i), i^* \in \text{neighborhood of } i$

$\sigma^2 = \text{variance of } MV(B_{pre}^i), i^* \in \text{neighborhood of } i$

식 (3)에서  $VD(B_{pre}^i)$ 는 블록  $B_{pre}^i$ 와 인접된 블록들에서 구한 동작 벡터 사이의 분산이 보상된 거리이다.  $VD(B_{pre}^i)$ 는  $MV(B_{pre}^i)$ 가 인접 블록들의 동작 벡터의 평균과 근접하고 분산이 작을 경우 작아진다. 즉, 인접한 블록들의 동작 벡터의 일관성이 높을 경우 고려되는 동작 벡터는 인접한 블록들의 동작 벡터의 평균과 근접하게 되므로  $VD(B_{pre}^i)$ 는 작아지지만  $CF(MV(B_{pre}^i))$ 는 커진다. 신뢰도가 높을수록 동작 벡터의 일관성이 높으므로 탐색 영역의 크기는 축소될 수 있다. 식 (3)에서  $K_1$ 과  $K_2$ 는 상수로서  $K_1$ 은 신뢰도가 취할 수 있는 최대값을 결정하고,  $K_2$ 는 신뢰도의 범위를 결정하는 역할을 한다.

정합 유사 함수는 블록 사이의 정합 척도로 사용되는 함수이다. 본 논문에서는 문자의 외관선 화소만 참여시켜 유사도를 구한다. 정합 유사 함수를 식 (4)와 같이

정의한다. 블록 정합의 기준은 최소 절대차(MAD : Minimum Absolute Difference)이며, 문자의 외곽선 화소들의 컬러의 최소 거리를 합에서 가장 작은 값의 위치를 찾는 것이다.

$$MAD(d_1, d_2) = \sum_{(x, y) \in R} (\sqrt{R_{dist}^2 + G_{dist}^2 + B_{dist}^2}) \quad (4)$$

$$\begin{cases} R_{dist} = |F_R(x, y, t) - F_R(x + d_1, y + d_2, t-1)| \\ G_{dist} = |F_G(x, y, t) - F_G(x + d_1, y + d_2, t-1)| \\ B_{dist} = |F_B(x, y, t) - F_B(x + d_1, y + d_2, t-1)| \end{cases}$$

여기서 R은 문자 구성요소 단위의 문자의 경계 화소, |d1|, |d2|는 기준 블록과 정합되는 블록 사이의 변위를 의미한다.  $F_R(x, y, t)$ 는 시간 t 프레임에 있는 좌표(x, y)의 화소 중 빨간색(Red) 값을 말한다.

#### 4.2.3 영역 보상 방법

장면 분할이 2회 이상 발생할 경우 참조할 원 영상이 없게 되며 따라서 확장적 블록 정합 방법을 사용할 수 없다. 이 문제를 해결 하기 위해서 본 논문에서는 영역 보상법을 제안하고자 한다. 복원될 영역은 자막 영역과 가장 근사한 값을 갖고 있는 주변 영역으로 볼 수 있다. 따라서 참조될 주변 영역을 어떻게 구하는가가 중요한 문제이다. 주변 영역을 구하는 방법은 그림 7과 같이 복원될 문자의 외곽선 상단과 하단에서 보상 될 영역을 추출한다. 자막 문자의 좌우 영역은 인접 문자의 접촉으로 보상 영역을 추출 할 수가 없다. 따라서 보상 영역을 문자 영역의 상단 및 하단 영역을 선택한다. 보상 영역의 폭은 자막의 폭과 동일하게 하고 일정 높이 이상으로 한다. 왜냐하면 높이가 2 화소 이하일 경우에는 질감을 표현할 수 없으므로 블록화 현상이 발생할 수 있다. 따라서 보상 영역의 높이는 3~5 화소 범위에서 결정한다. 영역 보상 방법은 그림 7와 같이 보상 영역을 반복적으로 보상하여 복원을 수행한다. 이것을 알고리즘화 하면 그림 8과 같다.

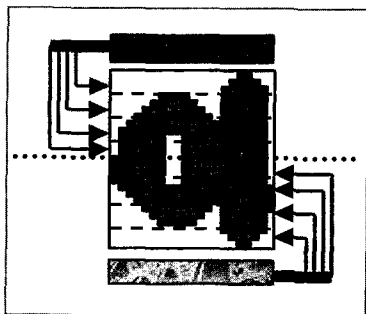


그림 7 영역 보상법

```
// 영역 보상 방법 //
===== 하향 방향 복원 =====
index = 0;
interval : 3 ~ 5;
st_y = block_start_y - (interval + 1);
while (incl < block_middle_y)
{
    incl = block_start_y + interval * index;
    inc2 = incl + interval * (index + 1);
    for (y = incl; y < inc2; y++)
        for (x = block_start_x; x < block_end_x; x++)
            image[y][x] = image[st_y + y - incl][x];
    index++;
}
===== 상향 방향 복원 =====
index = 0;
st_y = block_end_y + 1;
while (incl < block_end_y)
{
    incl = block_middle_y + interval * index;
    inc2 = incl + interval * (index + 1);
    for (y = incl; y < inc2; y++)
        for (x = block_start_x; x < block_end_x; x++)
            image[y][x] = image[st_y + y - incl][x];
    index++;
}
```

그림 8 영역 보상법의 알고리즘

## 5. 실험 및 결과 분석

실험 환경은 Windows NT상에서 750 MHz 속도를 갖는 PC상에서 구현하였다. Visual C++6.0 언어를 사용하였으며 비 압축 비디오 영상을 실험 영상으로 사용하였다. 비디오 영상은 RT5 Encoding 보드를 사용하여 file.mpg 파일을 획득한 후, 소프트웨어적으로 Decoding 하여 실험에 사용하였다.

실험에 사용된 영상은 영화 "매트릭스"와 "마야"를 실험에 사용하였다. 실험 결과 자막이 삽입된 곳은 대부분 대화 장면이나 설명을 위한 장면으로 물체의 움직임이나 카메라 동작이 크게 발생하지 않는 것을 볼 수 있었다.

그림 9는 기존의 블록 정합 방법과 본 논문에서 사용한 확장적 블록 정합 결과의 차이점을 보여주고 있다. 그림9(b)는 블록 정합 방법을 이용한 결과로 팔 부분에서 정합의 오류가 발생됨을 볼 수 있다. 그림9(c)는 확장적 블록 정합 방법을 이용한 결과를 보여 주고 있으며 기존의 방법보다 복원이 잘 됨을 볼 수 있다.

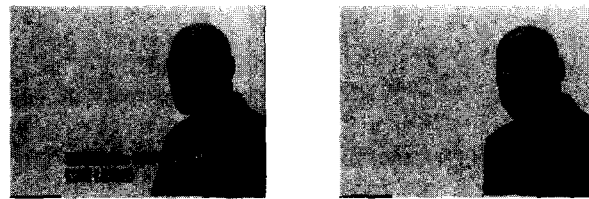
장면 전환이 2회 발생하여 참조할 원 영상이 없는 경우, 영역 보상법을 이용하여 복원된 결과를 그림 10에서 보여 주고 있다. 벡터 부분에서 약간의 끈김 현상이 발생됨을 볼 수 있다.

그림 11은 영화 "매트릭스" 중에서 장면 분할이 2회 발생한 부분의 복원한 결과를 보여 주고 있다. 그림 11



(a)원 영상 (b) 블록 정합 방법 (c) 확장적 블록 정합 방법

그림 9 확장적 블록정합 방법을 이용한 복원 결과



(a)원 영상 (b)영역 보상법을 적용한 결과

그림 10 영역 보상법을 이용한 복원 결과

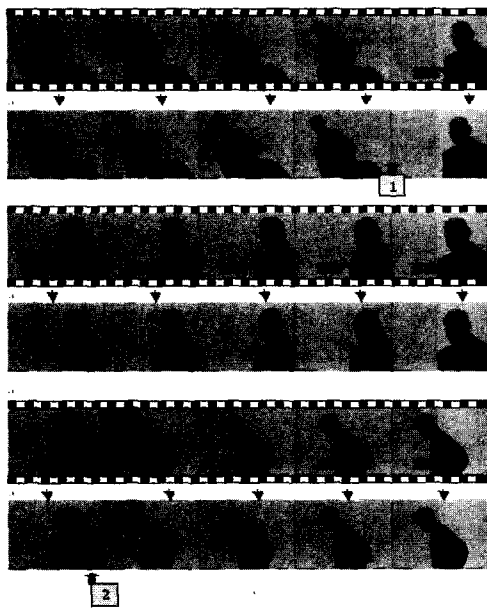


그림 11 장면 전환이 2회 발생한 영상의 복원 결과

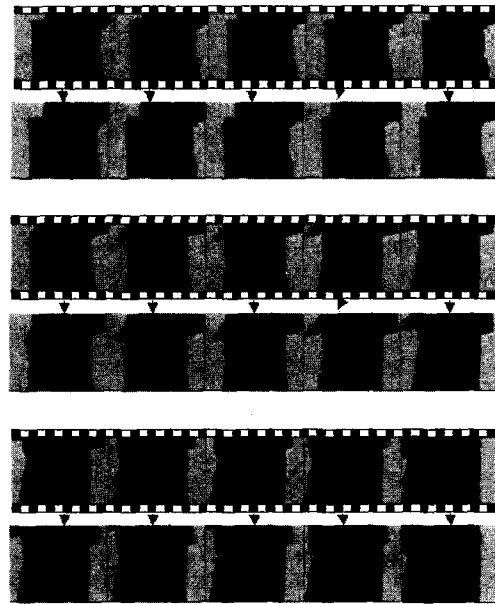


그림 12 영화 “매트릭스” 복원 결과

의 화살표는 장면 전환이 발생한 지점을 나타내 주고 있다. 첫 프레임부터 (1)번까지는 확장적 블록 정합 방법을 이용하여 복원한 결과를 보여 주고 있다. (1)번부터 (2)까지는 영역 보상법에 의한 복원 결과를 보여주고 있다. 마지막 프레임부터 (2)번까지는 역 방향으로 확장

적 블록 정합 방법을 이용하여 복원한 결과를 보여주고 있다. 실험결과 확장적 블록 정합 방법을 이용한 원영상 복원은 잘 수행됨을 볼 수 있었으며, 장면이 2회 이상 발생된 부분의 복원도 원만히 수행됨을 볼 수 있다.

그림 12는 “매트릭스”영화 중에서 일부를 복원한 결

과이다. 복원 결과가 좋음을 볼 수 있다.

## 6. 결론

본 논문에서는 확장적 블록 정합 방법과 영역 보상 기법을 이용하여 원 영상을 복원하였다. 확장적 블록 정합 방법의 사용으로 기존의 방법에서 발생할 수 있는 오 정합의 확률을 줄일 수 있었으며 복원의 정확을 기할 수 있었다.

또한 기존 방법에서 처리하지 못하였던 2회 이상의 장면 전환 문제, 즉 장면 전환 지점과 지점 사이의 복원 문제를 영역 보상법으로 처리함으로써 보다 완벽한 복원 시스템 구축할 수 있었다. 그러나 영역 보상 방법은 주위의 영역을 참조하여 복원하기 때문에 복원 시 약간의 부 자연스런 현상이 발생함을 볼 수 있다. 따라서 향후 연구 방향은 영역 보상법의 개선 또는 다른 접근 방법을 시도하여야 한다고 본다.

## 참 고 문 헌

- [1] 오원근, 유병문, 이인동 외2인, "영상처리 처리 기술을 이용한 조사당 벽화의 복원", 한국과학기술연구원/시스템공학연구소 연구보고서(E21131), 1992.
- [2] 전병태, 이재연, 배영래, "블록 정합을 이용한 비디오 자막 영역의 원 영상 복원 방법", 방송공학회논문지, Vol.5, No.1, pp.113~122, 2000.
- [3] 정기철, 남수진, 한정현, "Automatic Text Detection and Removal in Digital Videos", 컴퓨터비전 및 패턴인식 추계워크샵(CVPR'2001), pp.167~168, Nov., 2001.
- [4] Chun, Byung Tae, Younglae Bae, Tai-Yun Kim, "Recovering original images for video caption areas using camera motion and video information", Machine Graphics & Vision Journal, Vol.9, Nos.1/2, pp. 187~200, 2000.
- [5] 정기철, 남수진, 한정현, "Automatic Text detection and removal in digital videos," 2001년 컴퓨터 비전 및 패턴인식 연구회 추계 워크샵, pp.167-168, 연세대, 2001.
- [6] 이미숙, 설상훈, 양윤모, 이성환, "셋 경계 검출 및 카메라 동작 분석을 통한 비디오 자동 파싱 방법", 정보과학회논문지(B), Vol.25, No.8, pp.1249~1257, 1998.
- [7] J. S. Boreczky and L. A. Rowe, "Comparison of Video Shot Boundary Detection Techniques", IS&T/SPIE, Vol. 2670, pp.170-179, Feb. 1996.
- [8] 이재현, 장옥배, "움직임 벡터를 사용한 점진적 장면 전환 검출", 정보과학회논문지(C), 제3권, 제2호, pp.207-215, 1997.
- [9] 한규서, 이재연, 정세운, 배영래, "평균값을 이용한 통합형 카메라 동작 추출 시스템 개발", 한국정보처리학회 추계학술대회논문집(상), Vol.8, No.2, pp. 777-780, 2001.
- [10] J.L.Barron, et.all, "Systems and Experiment : Performance of Optical Flow Techniques", Int'l J. of Computer Vision, Vol. 12, 1994 pp.43-77
- [11] P.Jolly, H.K.Kim, "Efficient automatic analysis of camera work and microsegmentation of video using spatiotemporal images", Image Commu. Vol. 8, 1996, pp.295-307
- [12] Kim, Hae-Kwang, "Efficient automatic text location method and content-based indexing and structuring of video database", Journal of Visual Communication and Image Representation, Vol.7, No.4, pp.336-344, Dec. 1996.
- [13] Lienhart, Rainer and Frank Stuber, "Automatic text recognition in digital videos", Proc. of the SPIE, Image and Video Processing IV, Vol. SPIE2666, pp.180-188, San Jose, 1996.
- [14] 전병태, 배영래, 김태운, "일반화된 문자 및 비디오 자막 영역 추출 방법", 정보과학회논문지 : 소프트웨어 및 응용, Vol.27, No.6, 2000.
- [15] Furht, Borko, Joshua Greenberg and Raymond Westwater, *Motion Estimation Algorithm for Video Compression*, Kluwer Academic Press, 1997.
- [16] 장석우, "카메라의 동작을 보정한 장면전환 검출", 숭실대학교 박사학위논문, 2000.
- [17] Jang, Seok-Woo, Kyu-Jung Kim and Hyung-II Choi, "Accurate Estimation of Motion Vectors using Active Block Matching", Inter. Conference on RSCTC'2000, Spriger-Verlag in Lecture Notes, pp. 489 ~ 493, 2000.



전 병 태

1986년 2월 한남대학교 전산과(학사).  
1989년 8월 숭실대학교 전산과(석사).  
2001년 2월 고려대학교 컴퓨터학과(박사).  
1992년 5월 IR52 장영실 상 수상 (과기부 장관상). 1989년 9월 ~ 현재 한국전자통신연구원 · 컴퓨터 · 소프트웨어 연구소 선임연구원 재직중. 관심분야는 동영상 처리, 컴퓨터 비전, 멀티미디어 영상처리 등.



배 영 래

1976년 서울대학교 해양물리전공(이학사). 1986년 한양대학교 전자계산 전공(공학석사). 1992년 Univ. of Kent 전자공학 전공(공학석사). 1995년 Univ. of Kent 전자공학 전공(공학박사). 1976년~1979년 해군장교. 1979년 ~ 1980년 삼성. 1980년 ~ 2002년 2월 ETRI 컴·소 연구소 영상정보처리연구팀장. 2002년 3월 ~ 현재 충북과학대학 전자상거래학과 교수. 관심분야는 영상처리, 컴퓨터 비전, 병렬처리