

# 성대 영상에서 에너지를 이용한 관심 영역 추출 (Region-of-Interest Detection using the Energy from Vocal Fold Image)

김 엄 준 <sup>†</sup> 성 미 영 <sup>\*\*</sup>  
(Eom Jun Kim)(Mee Young Sung)

**요 약** 본 논문에서는 비테오스트로보키모그래피(Videostrobokymography) 시스템에서 영상중의 관심 영역을 추출하는 효율적인 방법을 소개하고자 한다. 비테오스트로보키모그래피는 성대 운동의 불규칙적인 움직임을 판단하여 자동으로 진단 파라미터를 구하는 의료 영상 시스템이다. 본 논문에서는 세 가지의 단계를 거쳐서 관심 영역을 추출하고 있다. 첫 번째로 최소 에너지를 이용하여 관심 영역의 중심이 되는 부분을 찾는다. 관심 영역 내에 있는 특징 점을 추출한 후 두 번째 단계로 한 라인(line) 영역에 대해 가로축을 따라서 평균값에 의한 에지를 선택한다. 최종 단계에서는 이 특징 값을 합병 알고리즘(merge algorithm)의 임계값으로 사용하여 관심 영역을 추출한다. 제안하는 알고리즘을 19명의 성대 영상에 적용하여 분석한 결과 성대를 촬영한 95%의 영상에서 관심 영역을 추출할 수 있었다. 본 연구에서 제안하는 관심 영역 추출 방법은 계산 량이 적어 200x280 크기의 영상을 초당 약 40프레임이상 처리하여 관심 영역을 추출할 수 있어 매우 효율적이다.

**Abstract** In this paper, we propose an effective method to detect the regions of interests in the Videostrobokymography System. Videostrobokymography system is a medical image processing system for extracting automatically the diagnosis parameters from the irregular vibratory movements of the vocal fold. We detect the regions of interests through three steps.

In the first step, we remove the noise in the input image and we find the minimum energy value in each frame. In the second step, we computed the edge by average value for the one line. In the third step, the regions of interests can be extracted by using the Merge Algorithm which uses the variance of luminance as the feature points.

We experimented this method for the vocal fold images of nineteen patients. In consequence, the regions of interests are detected in most vocal fold images. The method proposed in this study is efficient enough to extract the region of interests in the vocal fold images with the frame rate of 40 frames/second and the resolution of 200x280 pixels.

## 1. 서 론

의료 영상은 환자의 신체 부위를 촬영하여 분석함으로써 환자의 이상 유무를 판별하는 중요한 자료가 된다. 환자의 이상 유무를 판별하는 중요한 자료인 의료 영상을 분석하기 위해서 의료 영상 처리 분야는 발전하게 되었

다. 하지만 의료 영상은 촬영 당시의 배경과 기기의 환경에 따라 다양한 영상이 나올 수 있고, 이를 각 경우에 따라 다르게 처리할 수 있는 일반적인 시스템을 구축하는 것은 거의 불가능하기 때문에 각 분야별로 자신의 시스템에 알맞은 영상 처리 및 분석 시스템을 개발하여 사용하고 있다.

영상에서의 영역 추출을 하는 방법중 대부분은 에지(Edge), 히스토그램(Histogram) 분석 및 경계 값[1]에 의한 방법이다. 이러한 방법은 임계값에 의해서 영역을 추출하기 때문에 임계값이 잘못 되었을 경우에는 정확한 결과를 얻기 어렵다. 이러한 문제점을 해결하기 위하여 그동안 여러 가지의 모델과 방법이 발표되었다. 공간 특

· 본 연구는 한국과학재단 지정 인천대학교 멀티미디어연구센터의 지원에 의한 것입니다.

<sup>†</sup> 비 회 원 : 인천대학교 전자계산학과  
ejkim@isis.inchon.ac.kr

<sup>\*\*</sup> 중 심 화 원 : 인천대학교 전자계산학과 교수  
mysung@lion.inchon.ac.kr

논문접수 : 1998년 10월 16일

심사완료 : 2000년 6월 9일

값을 이용한 방법[2], 유전자 알고리즘을 이용한 방법[3], 움직임 벡터를 이용한 방법[4][12], 허프(Hough) 변환을 이용한 방법[5][13], 그리고 영상의 에너지를 이용한 방법[6] 등의 영상 추출을 위한 여러 가지 방법이 좋은 결과를 보이고 있다. 또한 최근들어 Active Contour Model(ACM)[7][14][15], Active Shape Model(ASM)[8] 등의 모델들이 발표되어 활발히 연구가 진행되고 있다.

공간 특징 값을 이용한 방법은 3x3의 윈도우를 선정하여 이들에서의 가우시안(Gaussian) 처리 값 및 확률적 계산을 하여 비교, 추출하는 방법이다. 이 방법에서는 객체와 배경을 포함한 영상에서 여러 개의  $k$ -집단( $k$ -population) 영상을 선정하여 분석하고 있다. 유전자 알고리즘을 이용한 방법은 생태계의 자연 유전과 자연 선택을 응용한 최적화(optimization) 알고리즘으로 방대하고 다양한 모델 함수에 대해 전역적인 최적화를 할 수 있는 장점을 가지고 있지만 계산 량이 크게 증가하고, 어느 정도 이상이 되면 인식률이 증가하지 않는 단점을 가지고 있다. 움직임 벡터를 이용한 방법은 세기를 기반으로 하여 점진적인 영상에 대한 변환 영역을 추출하고 있는데 이는 히스토그램의 차와 선형상관계수를 이용하여 영역을 검출하고 있다. 움직임 벡터를 이용한 방법은 다양한 형태의 탐색 물체에 대해서 동적으로 직선 성분 요소를 결정하여 자동적으로 임계치를 설정하는 장점을 가지고 있지만 곡선 성분이 많이 있는 탐색 물체에 대해 정확도가 낮은 단점을 가지고 있다. 영상의 에너지를 이용한 방법은 윤곽선을 최소 에너지 상태로 정의하고 있으며, 에너지 함수로는 스플라인(spline) 곡선을 사용하고 있다. 이 방법은 윤곽선의 초기 위치로부터 에너지를 최소화하는 과정을 통하여 스네이크(snake) 모델이 물체의 윤곽선으로 수렴하도록 하여 윤곽선을 구하고 있다. 이는 형태가 변하는 물체에 대해서 능동적으로 윤곽선을 추출할 수 있는 장점이 있지만 에너지의 최소화 방향을 찾을 수 있는 위치에 초기함수가 주어져야 하며, 물체의 이웃에 있는 잡음(noise)에 영향을 받을 수 있는 문제점을 가지고 있다. 이들 방법은 제한된 조건에서 효율적으로 추출해 낼 수는 있지만 모든 영상에 대해 일반적으로 적용되기는 어렵다. 이러한 방법들은 영상의 형태가 시간적으로 변형되지 않고, 밝기가 일정하게 유지될 수 있어야만 영상을 추출할 수 있다.

본 논문에서는 영상처리를 이용하여 후두 기능을 진단하기 위한 정량적 지표를 보여주는 비데오스트로보키모그래피(Videostrobokymography) 시스템[9]에서 성대의 관심 영역을 추출하는 방법을 보이고자 한다. 성대의 필립을 CCD카메라로 촬영한 영상은 비디오 테이프에 저장

된 후 영상 캡처 보드에 의해 그레이 영상으로 변환된 후 저장된다. CCD카메라에 의해서 촬영된 환자의 성대 움직임 영상은 촬영할 때마다 농도의 변화, 움직임 그리고 밝기 등의 영상 특징이 달라지게 된다. 이러한 동적인 특징의 변화는 관심 영역의 추출을 어렵게 만든다. 앞서 열거한 영역 추출 방법들은 동적인 성대 영상에 적용하기에는 에너지가 많고, 각 상황에 따라 특징 값이 달라지는 문제점, 그리고 곡선 성분이 많이 존재하는 문제점으로 인하여 적용하기 어렵다. 본 연구에서는 성대 영상에서 동적인 특징을 발견하는 방법과 동적인 변화를 보이는 관심 영역을 빠르고 정확하게 추출하는 방법을 제시한다. 본 연구에서는 다음의 세 가지 단계를 거쳐서 관심 영역을 추출하고 있다. 첫 번째로 최소 에너지를 이용하여 관심 영역의 중심이 되는 부분을 찾는다. 관심 영역 내에 있는 특징 점을 추출한 후, 두 번째 단계로 한 라인(line) 영역에 대해 가로축을 따라서 평균값에 의한 에너지를 선택한다. 최종 단계에서는 이 특징 값을 합병 알고리즘(merge algorithm)의 임계값으로 사용하여 관심 영역을 추출한다.

2장에서는 본 연구의 배경이 되었던 성대 운동 영상의 분석을 통해 파라미터를 구하는 비데오스트로보키모그래피에 대해서 간단하게 알아보고, 3장에서는 영역 추출의 각 단계별 방법론을 설명한다. 4장에서는 여러 가지의 실제 영상의 분석을 통한 평가를 하고, 5장에서는 결론을 내리고자 한다.

## 2. 비데오스트로보키모그래피(Videostrobokymography)[9]

비데오스트로보키모그래피는 CCD 카메라를 이용해서 초당 150~300번 열리고 닫히는 성대를 특수한 기법으로 고속 촬영하여 비디오 테이프에 저장한 후, 이 영상을 영상 캡처 보드를 통하여 컴퓨터로 초당 약 30 프레임씩 120프레임의 입력을 받아 그 영상을 처리하는 후두 기능 분석 기법이다. 이 방법은 사용자가 성대의 움직임을 직접 관찰하면서 시스템이 제공해주는 키모그램(kymogram)과 여러 가지 정량적인 지표들을 이용하여 성대의 여러 부분의 진동 양상을 비교, 평가할 수 있게 해준다.

비데오스트로보키모그래피 기법은 정량적 지표들을 추출할 때 수행해야 할 몇 가지 전 처리 과정들을 포함한다. 그 첫 번째는 관심 영역 추출 과정이다. 이 과정은 각 영상이 CCD 카메라의 성능 및 위치에 따라서 영상의 값이 달라질 수 있는데 이를 수정하기 위한 것이다. 두 번째는 기울기 보정 과정이다. CCD 카메라에 의해서 촬

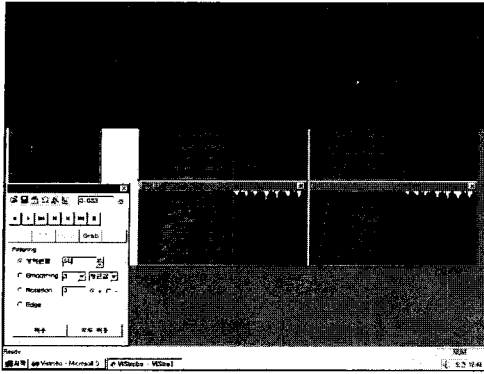


그림 1 kymogram

영한 영상은 기울어질 확률이 많기 때문이다. 세 번째는 위치 보정 과정이다. 이는 CCD 카메라나 환자의 움직임에 의해서 각 프레임의 관심 영역의 위치가 변할 수 있으므로 이를 수정하는 과정이다. 모든 전 처리 과정들이 최종 결과를 내기 위해서 중요하지만 이 중에서 관심 영역 추출이 첫 번째로 이루어져야만 한다. 이 과정이 정확히 이루어져야만 그 이후의 과정이 명확하게 처리될 수 있다. 그림 1은 비데오스트로보키모그래피에 의해서 구하여진 최종 키모그램을 보여주고 있다. 이 그림은 관심 영역을 추출하지 않고 수행하였으며, 4개의 키모그램에서 가운데의 검은 점 같은 부분은 성대가 열려있는 부분이며, 그 나머지 부분은 성대가 닫혀 있는 부분이다. 이러한 키모그램을 정확히 분석하기 위해서는 반드시 관심 영역에 대한 추출 및 분석이 행해져야 한다.

3. 영상 특징 및 추출 알고리즘

3.1 성대 영상의 특징

관심 영역을 추출하기 위해서는 영상의 특징을 아는 것이 중요하다. 이 특징에 따라서 어떤 방법으로 어떤 알고리즘을 써서 추출할 것인가가 결정되기 때문이다. 앞서 설명한 바와 같이 각 시스템에 알맞은 방법을 사용하여 영역 추출을 하여야만 한다. CCD카메라로 성대의 움직임을 촬영한 영상에서 추출하고자 하는 관심 영역은 그림 2에서 가운데의 검은 부분이다. 추출하기 위한 영역은 그림 2와 같이 주위의 영상보다 밝기값이 낮고, 또한 세로축이 가로축보다 긴 타원의 모양을 하고 있다. 성대 영상은 그때 그때마다의 환경에 따라서 여러 가지의 특징이 달라질 수 있다. 여기에서는 공통으로 가지는 특징을 설명한다.

성대 영상에서의 관심 영역을 추출하기 위한 영상의 특징은 다음과 같다.

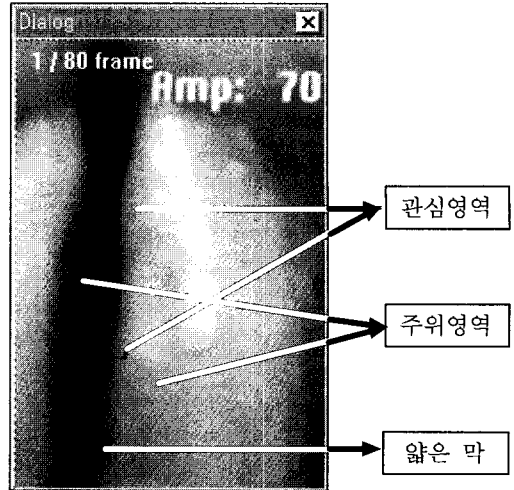
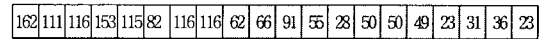


그림 2 성대 영상

- 1) 성대가 떨리고 있는 부분은 떨리고 있지 않은 부분보다 밝기값이 낮다.
- 2) 관심 영역의 밝기값은 그림 3과 같이 가장 낮은 값을 가지는 지점으로부터 그 주위로 점차적으로 값이 높아진다.



<-- 주위 영역 | 관심 영역 -->



그림 3 한 선의 밝기값

- 3) 관심 영역은 정상일 경우 세로축이 가로축보다 긴 하나의 타원 모양을 한다.
- 4) 관심 영역은 성대의 떨림에 의해서 아래에서부터 출발하여 위로 움직인다.
- 5) 성대의 떨림은 정상일 경우 세로축을 중심으로 좌우 대칭인 모양을 하고 있다.
- 6) 관심 영역은 비정상일 경우 아래, 위 두 부분으로 나누어 질 수 있다.

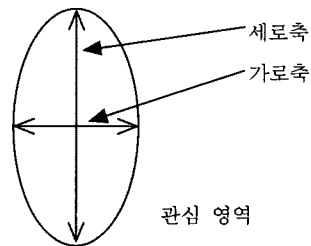


그림 4 성대의 모양

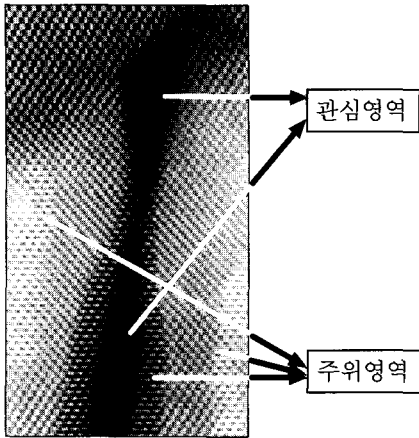


그림 5 아래, 위로 분리된 성대의 관심 영역

7) 성대가 떨리고 있는 영상에는 안쪽에 얇은 막이 있을 수 있으며, 이는 열려있는 것으로 간주하지 않는다.

위와 같은 특징에 의해 성대 영상에서 관심 영역을 추출하는 방법을 알아본다. 보편적인 관심 영역 추출 방법은 에지를 구하는 것이다. 그러나 성대 영상에서 일반적인 에지 추출 방법을 적용하는 경우 문제점이 있다. 성대 영상은 관심 영역 이외의 곳에도 에지가 존재할 수 있고, 관심 영역 안에도 에지가 될 수 있는 곳이 여러개 존재할 수 있다. 또한 CCD카메라에 의해서 촬영된 영상은 촬영할 때마다 환경의 영향으로 인해서 연속적으로 촬영된 영상이라 할지라도 각각 특징이 다르다. 그러므로 각각 다른 특징을 가지는 영상마다 서로 다른 경계 값을 가져야 한다. 각 영상에서의 경계값을 추출하기 위해서 각 프레임마다 먼저 관심 영역 안에 있는 최소 에너지의 위치를 구한다. 이 위치에 대하여 각 프레임간에 서로 보정을 하고 나서 그 위치로부터 출발하여 가로축으로 밝기값의 차분을 계산하고 차분이 최대인 값을 구한다. 이 값을 경계값으로 하여 최소 에너지 지점으로부터 병합 알고리즘을 적용하여 관심 영역을 추출하게 된다.

### 3.2 에너지 최소화에 의한 특징점 추출

#### 3.2.1 각 프레임의 후보 특징점 추출

본 연구에서는 이차원 영상을 사용하였으며, 이차원 영상의 에너지는 주변 화소 사이의 차이에 해당하는 값이다. 그러므로 전체 촬영한 성대 영상에서 관심 영역을 추출하기 위해 영상의 에너지를 이용하여 각각의 프레임에서 특징점을 구한다. 각 프레임에서 관심영역에 해당하는 특징점들을 추출하는 과정은 다음과 같다.

1. 한 프레임에서 한 줄씩 각 줄에 대한 최소 에너지의 후보점들을 추출한다. 각 줄에서 추출한 후보점들은 오

목하게 들어간 부분이며, 여러 개가 존재할 수 있다.

2. 한 프레임에서 구해진 후보점들의 전체 평균 및 표준편차를 구한다. 그리고, 각 줄에서 구해진 후보점이 전체 평균으로부터 표준편차 이내에 해당하는 값인지를 비교한다. 비교하는 후보점이 평균에서부터 편차 이내에 해당하는 값일 경우 그 줄에서의 후보점으로 취한다.

3. 각 줄에서 구해진 후보점들 중에서 최소 에너지의 값과 최소 에너지를 가지는 좌표를 취하여 후보 특징점으로 선택한다.

한 프레임 영상의 전체 크기는  $n \times m$ 이고, 각 좌표  $(x_n, y_m)$ 에서의 밝기값을  $P_{nm}$ , 각 지점에서의 에너지는  $E_k$  ( $0 \leq k \leq n \times m$ )이라 한다. 한 프레임의 각 줄마다 최소 에너지 및 좌표를 구한다.

$$E_{ij} = P_{i(j+1)} - P_{i(j-1)}, \quad 0 \leq i \leq n, 1 \leq j \leq m-1 \quad (1)$$

$P_{i(j+1)}$  : 좌표  $(i, j+1)$ 에서의 밝기값

$P_{i(j-1)}$  : 좌표  $(i, j-1)$ 에서의 밝기값

이를 행렬로 나타내면 식(2)와 같다.

$$E_{ij} = \begin{bmatrix} E_{00} & E_{01} & E_{02} & \dots & E_{0j} \\ E_{10} & E_{11} & E_{12} & \dots & E_{1j} \\ E_{20} & E_{21} & E_{22} & \dots & E_{2j} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ E_{n0} & E_{n1} & E_{n2} & \dots & E_{nj} \end{bmatrix} \quad (2)$$

식(2)에서 각 열마다 음수에서 양수로 변하는 지점 즉, 오목한 지점을 구하고, 그 중에서 최소 에너지를 구한다.

$$\begin{aligned} E_{\min 0} &= \text{MIN}(E_{00}, E_{01}, E_{02}, \dots, E_{0j}) \\ E_{\min 1} &= \text{MIN}(E_{10}, E_{11}, E_{12}, \dots, E_{1j}) \\ E_{\min 2} &= \text{MIN}(E_{20}, E_{21}, E_{22}, \dots, E_{2j}) \\ &\vdots \\ E_{\min i} &= \text{MIN}(E_{i0}, E_{i1}, E_{i2}, \dots, E_{ij}) \end{aligned} \quad (3)$$

식(3)에 의해 각 행에서의 최소 에너지를 구할 때 그의 좌표  $Q_0, Q_1, Q_2, \dots, Q_i$ 를 식(4)와 같이 구한다.

$$\begin{aligned} Q_0 &= E_{\min 0} \text{의 좌표값 } j \\ Q_1 &= E_{\min 1} \text{의 좌표값 } j \\ Q_2 &= E_{\min 2} \text{의 좌표값 } j \\ &\vdots \\ Q_i &= E_{\min i} \text{의 좌표값 } j \end{aligned} \quad (4)$$

식(3)과 식(4)에 의해서 구해진  $E_{\min 0}, E_{\min 1}, E_{\min 2}, \dots, E_{\min i}$ 의 최소 에너지를 원소로 하는 집합  $A$ 와  $Q_0, Q_1, Q_2, \dots, Q_i$ 의 최소 에너지 좌표를 원소로 하는 집합  $B$ 는 식(5)와 같다.

$$\begin{aligned} A_{\text{frame}} &= \{x|x \text{는 한 프레임의 각 행에서의 최소 에너지}\} \\ B_{\text{frame}} &= \{y|y \text{는 한 프레임의 각 행에서의 최소 에너지 좌표}\} \end{aligned} \quad (5)$$

$1 \leq \text{frame} \leq$  전체 프레임의 수

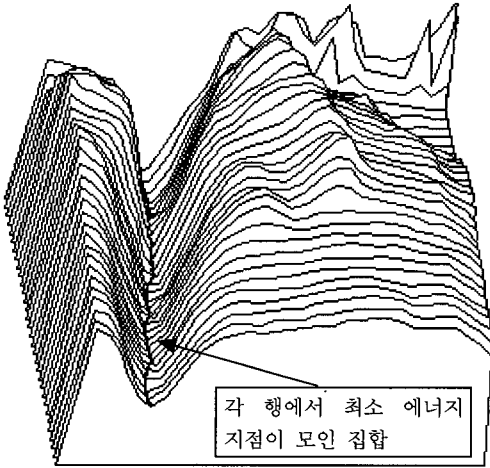


그림 6 그림 2의 Luminance Surface

그림 6은 그림 2의 상대 영상을 luminance surface로 나타낸 그림이다. 이 그림에서 계속 부분의 검은 선은 각각의 열마다 구하고자 하는 좌표를 연결하여 얻은 선이다. 각각의 프레임에 대하여 식(5)에서 구한 집합 A의 원소 중 가장 최소값을 취하면 각 프레임에서의 후보 특징점이 된다.

$$C = \{x_n \text{는 각 프레임에 해당되는 집합 } A \text{에서의 최소값}\} \quad (6)$$

$$D = \{y_n \text{는 집합 } C \text{에서의 각 원소에 대한 좌표값}\}$$

식(6)에서의 집합 C는 각 프레임에 해당되는 집합 A에서의 최소값을 원소로 하며, 집합 D는 집합 C에 구해진 각 원소와 대응되는 집합 B에서의 원소값을 원소로 한다.

그림 7은 집합 C에서의 원소들 중 해당 프레임 원소의 위치를 영상에 흰 점으로 표시한 그림이다. 현재 80개의 영상 프레임 중 첫 번째 프레임을 보여주고 있으며,

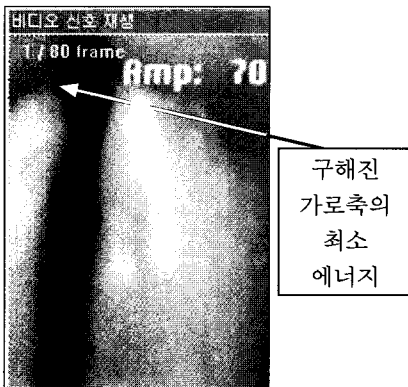


그림 7 한 프레임에서의 후보 특징점

식(1)부터 식(5)까지의 계산 결과에 의해 한 프레임의 각 행에서의 최소 에너지를 구한 집합 A와 그에 해당하는 좌표의 집합 B가 구해진다. 이 후, 집합 A에서의 최소값을 구하게 되는데, 그림 7에서는 두 개의 동일한 값을 가진 최소 에너지 지점이 구해지고 있다. 집합 C에는 두 개의 동일한 값을 가진 원소가 존재하게 되며, 이에 해당하는 좌표의 집합인 D에는 다른 값을 원소로 가지게 된다.

3.2.2 각 프레임의 특징점 보상

각 프레임마다 최소 에너지에 의해 추출된 후보 특징점들은 서로 다른 위치에 존재하고 있다. 이는 촬영대상이 움직이거나 카메라가 이동하는 등의 순간의 상황 변화에 의해서 입력된 영상의 값이 달라졌기 때문이다. 즉, 프레임마다 밝기가 달라질 수 있고, 움직이는 영상일 경우 좌우나 아래위로 그림 8과 같이 관심 영역이 이동하게 된다. 이는 성대 운동의 분석을 위한 최종 파라미터에 왜곡을 가져오게 된다.

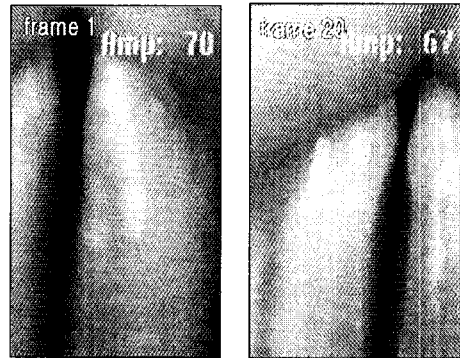


그림 8 프레임 1과 프레임 20 영상에서의 관심 영역의 이동

관심 영역의 이동에 대한 보상을 하기 위하여 프레임마다 구해진 각 후보 점들의 최소 에너지를 가지고 전체 평균과 표준편차를 구한다. 이렇게 계산하는 이유는 관심 영역이 움직이는 영상이며, 움직임 보상을 하기 이전에 각각의 프레임에서 구해진 후보 특징값이 잘못 구해진 값인지를 판별하기 위해서이다. 각 후보 점의 편차의 절대값이 표준편차보다 높으면 이는 잘못 구해진 것으로 간주하고 그 프레임의 그 다음 최소 에너지 지점에 대해서 다시 계산한다. 이의 반복에 의해 각 프레임에서 관심 영역에 있는 최소 에너지 지점을 찾는다. 최소 에너지 지점의 값들을  $x_1, x_2, \dots, x_n$ , 평균을  $m$ 이라 하면, 편차  $(x_i - m)$ 와 표준편차를  $\sigma$ 를 다음과 같이 구할 수 있다.

$$\sigma^2 = \frac{1}{2} \sum_{i=1}^n (x_i - m)^2, \sigma = \sqrt{\sigma^2} \quad (7)$$

$$(x_i - m) < \sigma : \text{허용} \quad (8)$$

식(7)과 식(8)의 계산에 의해서 구해진 각 프레임의 좌표들에 대해서 허용된 맨 첫 프레임의  $x, y$  축 값을 기준으로 나머지 프레임에 이동시킨다. 성문이 열려 있는 부분만을 본다면 위에서 설명한 것과 같이 마름모의 형태를 갖고 있으며, 움직임의 방향은 일정하다. 각 마름모의 위아래 꼭지점을 연결한 선들이  $y$  축과 나란하도록 세로축선들을 이동시켜 특징점의 움직임을 보정한다.

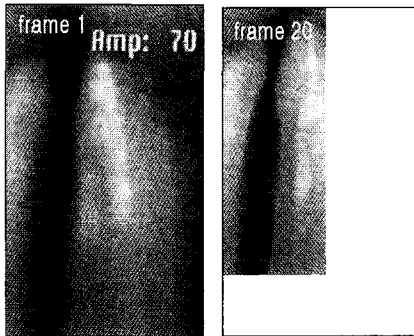


그림 9 각 프레임에서의 특징점 보상

### 3.3 관심 영역 추출

#### 3.3.1 에지 추출

각 프레임마다 최소값의 좌표에 대한 분산을 구하여 다른 값을 가지는 프레임의 최소값과 좌표를 보정한다. 구해진 최소값 좌표를 중심으로 가로축을 따라 마스크를 취하고, 그 평균값을 구해나간다. 구해진 평균값과 이동 후에 마스크를 취한값을 비교하여 만일 마스크를 취한값이 평균값보다 작으면 이를 허용하여 다시 평균값을 계산한다. 반대로 에지가 평균값보다 크면 최종 에지로 간주한다. 그림 10은 그림 7에서 구한 최소 에너지 위치에

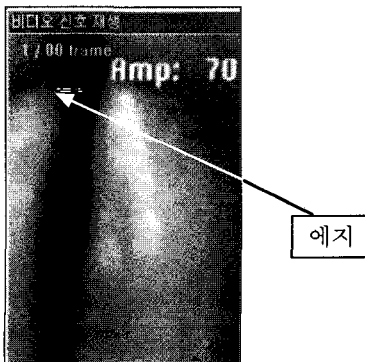


그림 10 그림 7의 에지를 구한 영상

서 출발하여 왼쪽으로 이동하면서 마스크를 취하여 구하여진 에지를 보여주고 있다.

에지를 구하기 위한 마스크(mask)는 아래와 같다.

-3	-1	0	1	3
----	----	---	---	---

#### 3.3.2 합병 알고리즘(Merge Algorithm)[10][17]

대부분 영상에서 같은 이미지 영역 안에 근접한 곳에 있는 각 픽셀(pixel)들은 비슷한 속성을 가지고 있다. 이러한 속성을 이용하여 영상에서의 영역 분할을 하는 방법으로 합병 알고리즘이 사용되고 있다. 이 알고리즘은 각 영역에서 주위에 있는 비슷한 속성을 가진 영역끼리 병합하여 영역을 확장해 나가는 방식이다. 즉, 최소 에너지 지점에서부터 출발하여 임계값에 알맞은 영역을 포함해가면서 확장하고 있다.

각 픽셀을  $s_i(i=1, \dots, N)$ 라하고, 픽셀들의 집합을  $P$ 라하고, 각각의 영상 영역을  $R_i$ 라 하면,  $R_i$ 에는 주위에 있는 비슷한 속성을 가진 픽셀들이 모여있는 집합이 된다. 각 집합에 대한 특징값으로 평균 및 표준편차를 구하고, 주위 영역의 특징값이 비슷한 영역끼리 서로 병합하여 다시 특징값을 구해 나가게 된다. 최초 식(6)에서 구한 최소 에너지 지점에서부터 시작하여 주변 영역으로 확장하기 위해서 합병 알고리즘(merge algorithm)을 사용하는데, 이는 여러 번의 단계를 통하여 영역 확장이 이루어진다. 각 단계를  $k$ , 각 단계의 영역을  $R_i(k)(i=1, \dots, N)$ 라 한다.

영역  $R_i(k)(i=1, 2, \dots, N)$ 는 초기에 구해진 최소 에너지의 값, 그리고 그의 좌표를 가진다. 먼저 하나의 위치  $X$ 마다 8개의 주변 영역의 값을 비교하여 비슷한 값을 가지는 지점을 포함시킨다. 최종적으로  $R_i$ 값이 구해지면 이에 대해서 평균과 분산을 구하여 각 영역마다의 특성 값으로서 식(9)와 식(10)과 같이 평균 및 표준편차를 구한다.

$$m_i = \frac{1}{n} \sum_{(k,l) \in R_i} P(k, l) \quad (9)$$

$$\sigma_i = \sqrt{\frac{1}{n} \sum_{(k,l) \in R_i} (P(k, l) - m_i)^2} \quad (10)$$

$P(k, l)$ 은 한 영역 안에 있는 각 지점의 값을 말한다. 각 영역은 자신의 영역 측정값과 주위 영역의 측정값을 비교하여 식(11)의 조건에 의해서 병합 여부를 결정한다.

$$|m_1 - m_2| < k\sigma_i \quad (i=1, 2), \quad k \text{는 상수} \quad (11)$$

식(11)에서 상수  $k$ 는 병합의 범위를 결정지을 수 있으며,  $k$ 가 작으면 병합은 세분화되고, 크면 병합은 포괄적으로 이루어진다.

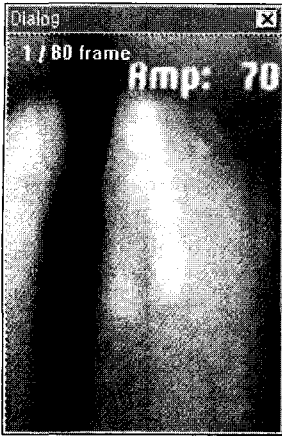
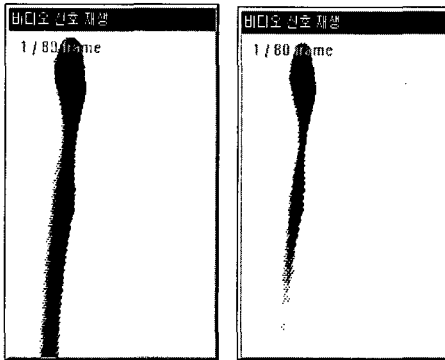


그림 11 합병 알고리즘의 적용 결과

음과 같다. 3.3.3절의 과정을 거친 후에 나온 결과의 영상에 식(6)에서 구해진 각 프레임에서의 후보 특징점의 집합 C와 D의 원소를 가지고 3.3.1절의 에지 추출 과정을 두 번째로 적용하여 에지와 에지에 해당하는 좌표를 구한다. 이 때 구해진 에지는 그림 12의 (a)에서 보여지는 검은 영상 범위 내에서 구하기 때문에 처음에 구한 에지보다 안으로 옮겨진 지점에 구해진다. 그러므로 처음에 구해진 에지와 두 번째 에지 추출과정을 적용하여 나온 에지와의 차이에 해당하는 영상이 얇은 막에 해당한다. 이와 같이 구해진 에지에 3.3.2의 합병 알고리즘을 적용한 결과 영상에서 관심 영역은 검은색으로 그 이외의 영역은 흰색으로 나타내면, 그림 12의 (b)와 같은 영상이 구해지며, 이는 구하고자 하는 영상의 최종 결과이다.



(a) 얇은 막을 포함한 영상 (b) 얇은 막을 제외한 영상

그림 12 관심 영역 추출

#### 4. 분석 및 평가

본 논문에서 제시한 성대 영상에서의 관심 영역 추출을 위한 알고리즘을 분석하기 위해서 실제 여러 환자의 데이터에 적용해 보았다. 또한 최소 에너지를 이용하여 윤곽선을 추출하는 알고리즘인 스네이크 모델을 적용하였을 경우와 비교해 보았다. 본 실험에서 적용하여본 스네이크 알고리즘은 변화하는 동적인 영상에서 관심 영역을 추출하는데 좋은 성능을 여러 논문 및 실험에서 보여주고 있다. 본 논문에서 실험한 영상은 비디오 테이프로부터 초당 30프레임으로 입력되는 동적인 영상이며 각각의 프레임에서 관심 영역을 추출하여야 하기 때문에 스네이크 알고리즘의 실험에 알맞은 영상이라 여겨지고, 스네이크 알고리즘에 본 실험에서 사용된 영상을 적용해 보았다.

#### 3.3.3 관심 영역 추출

에지 추출에서 구해진 경계값을 합병 알고리즘에 적용하여 최종 관심 영역을 추출한다. 그림 11은 3.3.2에서 설명한 합병 알고리즘을 그림 2에 적용한 그림이다.

그림 12의 (a)는 위의 알고리즘에 의해서 나온 결과를 관심 영역은 검은 색, 그 외의 영역은 흰색으로 구분한 결과이다.

현재까지 구해진 영상인 그림 12의 (a)는 얇은 막을 포함한 영상으로서 최종 결과가 아니다. CCD 카메라에 의해 촬영된 성대 영상에는 성대의 고속 움직임에 의한 잔상이 있다. 이 막은 성대가 닫힌 부분이며, 이러한 얇은 막까지 추출하기 위해 최소 에너지 지점에서부터 출발하여 에지 검출, 합병 알고리즘을 그림 11에 다시 한번 적용하면 그림 12의 (b)를 추출할 수 있다.

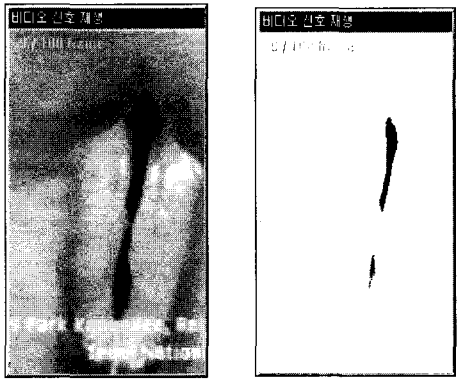
그림 11에서 얇은 막을 제외하는 과정을 설명하면 다

서론에서 설명한 바와 같이 CCD카메라에 의해서 촬영을 하여 비디오 테이프에 저장된 환자의 성대 영상을 그레이 모드로 입력받았다. 시스템은 인텔 펜티엄 200MHz, 메모리는 32Mbyte, 최소 200Mbyte가 남아 있는 HDD, 개발툴은 Visual C++ 4.0등의 환경에서 분석을 하였다. 평가를 하기 위해 증점적으로 살펴본 사항은 밝기가 다르게 촬영된 각 환자의 성대 영상에 대해서 본 알고리즘이 성대의 열린 부분을 추출해 낼 수 있는가와 얇은 막이 포함되어 있는 영상에서 얇은 막을 제거한 진짜 관심 영역을 추출해 낼 수 있는가이다. 결론적으로 19명의 환자 성대 영상을 분석한 결과 1274개의 영상 중 관심 영역의 농도가 관심 영역 이외의 농도와 비슷한 경우를 제외한 95%의 영상에서 정확히 추출할 수 있었다.

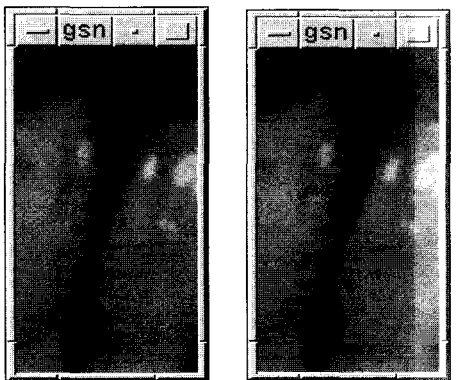
평가를 하기 위해서 에너지 최적화에 의한 관심 영역을 추출하는 스네이크 알고리즘의 결과와 비교해 보았다 [6]. 스네이크 알고리즘은 추출하고자 하는 영상의 위치

를 선택한 후 실행을 하게 되면 선택한 각 지점에서부터 출발하여 최소 에너지 지점을 찾아가게 되는데, 예지 구별이 확실한 영상에서는 관심 영역을 추출할 수 있었다. 스네이크 알고리즘은 SUN Sparc Solaris 2.5.1 시스템에서 싱가포르의 Nanyang Technological University(NTU)의 GSNAKE API ver 1.0[18]을 실행하였다.

4.1절에서부터 4.6절까지는 성대 영상의 여러 가지 예를 들고, 각각의 성대 영상에 본 논문에서 제안한 방법을 적용한 결과와 스네이크 알고리즘을 적용한 결과를 비교하였다. 그림 13부터 그림 18까지의 그림 중에 그림(a)는 원래 영상이며, 그림(b)는 본 논문에서 제안한 방법을 적용한 결과이다. 그림(c)는 원래 영상에 추출하고자 하는 영역의 위치를 선택한 그림이며, 그림(d)는 그림(c)에 스네이크 알고리즘을 적용한 결과이다.



(a) 최초영상 (b) 본 알고리즘에 의해 구해진 영상



(c) 스네이크 알고리즘을 위해 초기 위치를 선택한 영상 (d) 스네이크 알고리즘에 의해 구해진 영상

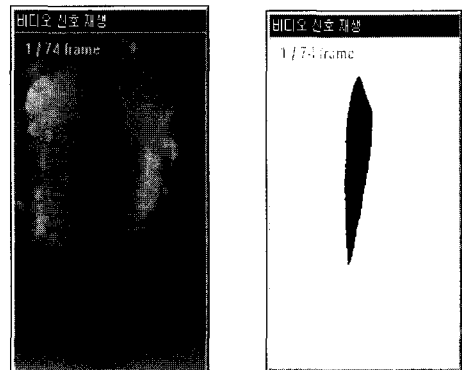
그림 13 밝게 촬영된 영상

#### 4.1 밝게 촬영된 영상

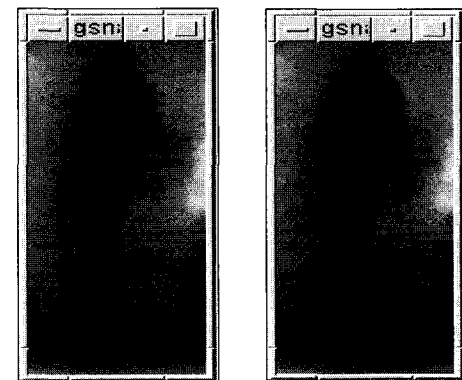
그림 13의 (a)는 본 논문에서 제안한 방법을 적용하여 관심 영역을 추출하기 위한 가장 이상적인 영상이다. 성대가 열린 부분이 상대적으로 어둡기 때문에 얇은 막을 제외한 관심 영역은 (b)와 같이 잘 추출되었다. 스네이크 알고리즘에서도 그림 13의 (d)에서 보는 바와 같이 어느 정도 관심 영역을 추출하였으나 미세하게 분리되는 곳은 추출하지 못하였다.

#### 4.2 성대 주위의 외곽 부분이 어두울 경우

그림 14의 영상은 성대 주위의 외곽 부분이 어둡게 촬영이 된 영상이다. 이 영상에서 추출하고자 하는 관심 영역이 관심 영역을 둘러싼 부분보다 상대적으로 어둡다. 그러므로 본 논문에서 제안한 방법에 의해서 추출되었다. 스네이크 알고리즘을 적용한 영상도 또한 잘 추출하였다.



(a) 최초영상 (b) 본 알고리즘에 의해 구해진 영상



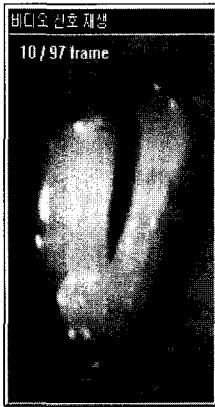
(c) 스네이크 알고리즘을 위해 초기 위치를 선택한 영상 (d) 스네이크 알고리즘에 의해 구해진 영상

그림 14 성대 주위의 외곽 부분이 어두울 경우

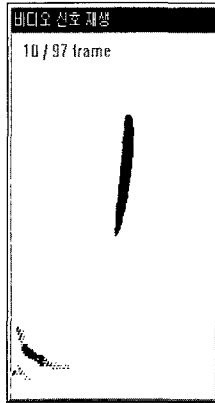


### 4.3 영상에서 관심 영역이 작은 경우

그림 15의 영상에서도 관심 영역을 잘 추출하였으나 관심 영역 이외의 주위 영역 중에 관심 영역 만한 밝기의 영역이 존재할 경우 그 부분까지 추출하고 있다. 그러나 스네이크 알고리즘으로는 이 영상에서 결과를 낼 수 없었다.



(a) 최초영상



(b) 본 알고리즘에 의해 구해진 영상



(c) 스네이크 알고리즘을 위해 초기 위치를 선택한 영상

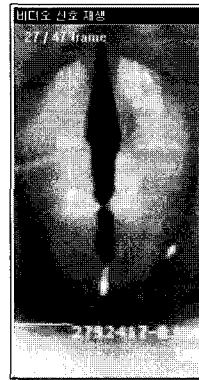


(d) 스네이크 알고리즘에 의해 구해진 영상

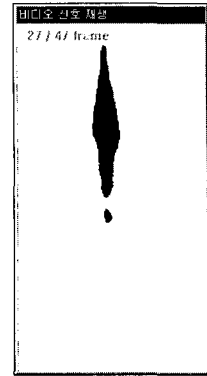
그림 15 영상에서 관심 영역이 작은 경우

### 4.4 관심 영역이 여러 개로 분리되어 존재할 경우

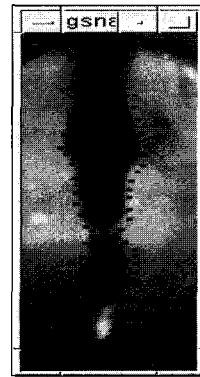
그림 16의 영상은 환자의 성대 이상에 의해서 관심 영역이 둘 이상으로 분리되어 있는 경우이다. 이 경우에도 분리되어 있는 관심 영역을 잘 추출하였다. 이 영상에서 안에 있는 얇은 막까지 제외되는 것을 알 수 있다. 스네이크 알고리즘에서는 확실히 분리되는 부만을 찾는 것



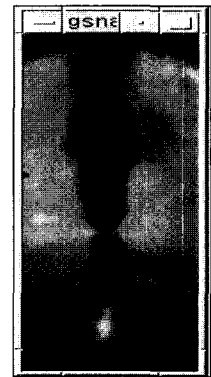
(a) 최초영상



(b) 본 알고리즘에 의해 구해진 영상



(c) 스네이크 알고리즘을 위해 초기 위치를 선택한 영상



(d) 스네이크 알고리즘에 의해 구해진 영상

그림 16 관심 영역이 여러 개로 분리되어 존재할 경우

을 볼 수 있다. 이 때, 얇은 막이 있는 부분의 관심 영역은 추출하지 못한다.

### 4.5 성대에 이상이 있는 환자의 성대 영상

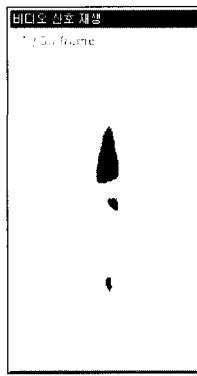
그림 17의 영상에서는 환자의 성대에 혹이 있는 것을 볼 수 있다. 성대의 떨림에 의해서 다른 부위는 열리고 닫히는 운동을 하는데 반하여 이 혹이 있는 부분은 계속 닫힌 상태로 존재한다. 이러한 환자의 이상 부위가 존재하는 영상에 대해서 성대의 관심 영역을 잘 추출하였다.

### 4.6 관심 영역의 주위가 어두울 경우

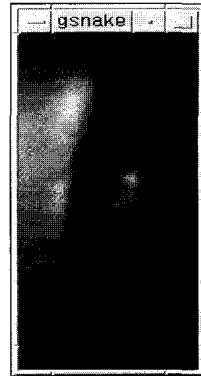
그림 18의 영상에서는 관심 영역에 대한 특징값을 찾아내지 못함으로써 성대의 관심 영역을 추출하지 못하였고, 스네이크 알고리즘에서는 수작업으로 관심 영역을 선택하고 있으므로 상대적으로 관심 영역을 추출할 가능성이 있다.



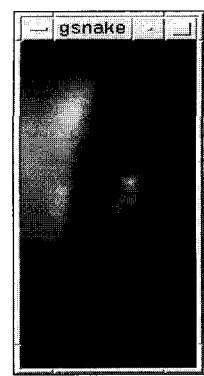
(a) 최초영상



(b) 본 알고리즘에 의해 구해진 영상

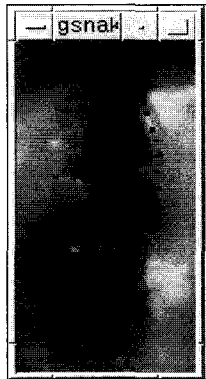


(c) 스네이크 알고리즘을 위해 초기 위치를 선택한 영상

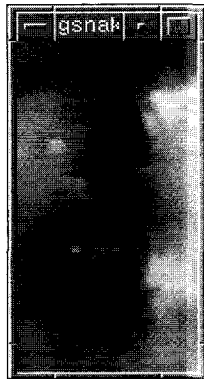


(d) 스네이크 알고리즘에 의해 구해진 영상

그림 18 관심 영역의 주위가 어두울 경우

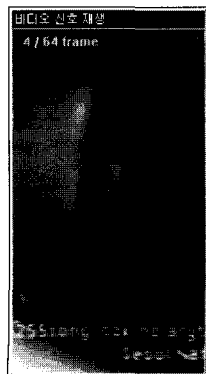


(c) 스네이크 알고리즘을 위해 초기 위치를 선택한 영상

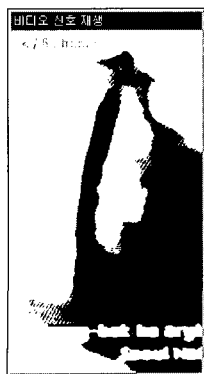


(d) 스네이크 알고리즘에 의해 구해진 영상

그림 17 성대에 이상이 있는 환자의 성대 영상



(a) 최초영상



(b) 본 알고리즘에 의해 구해진 영상

### 5. 결론

지금까지 성대 영상에서 관심 영역을 추출하기 위하여 성대 영상에 대한 특징과 문제점을 소개하였고 이에 대한 해결책을 제시하였다. 움직이는 성대를 촬영한 영상에서 각 프레임마다 변하는 영상의 특징을 찾아내기란 매우 어렵다. 본 논문에서는 동적으로 변화하는 성대 영상에서 동적으로 변하는 관심 영역을 추출하는 기법을 제안하였으며 이 기법은 아래의 두가지로 요약될 수 있다.

- 최소 에너지에 의한 특징값의 추출
- 위에서 구한 특징값에 의한 분할-합병 알고리즘의 적용

에지 추출과 영상의 에너지에 의해서 각 프레임에서의 특징값을 구한 후 분할-합병 알고리즘을 적용함으로써 성대의 관심 영역을 추출하였다. 여기에서 제안하는 영역추출 방법은 동적인 변화를 보이는 영상에서 관심 영역을 추출할 수 있을 뿐 아니라 계산량이 적어 200x280의 영상을 대상으로 초당 약 40프레임의 관심 영역을 추출할 수 있다. 또한 제안하는 알고리즘의 성능을 분석하기 위해 최소 에너지에 의해 동적인 변화를 가지는 영상에서 관심 영역을 추출할 수 있는 알고리즘인 스네이크 알고리즘과 비교해 보았다. 그 결과 스네이크 방식으로는 초기에 선정된 곡선에 의해서 많은 영향을 받을 뿐만 아니라 파라미터 값을 각 프레임마다 결정을 해주었음에도 불구하고 이에 의한 만족할 만한 결과를 얻지 못하였다. 성대 영상은 다양한 모양의 곡선을 가지고 있는데 스네이크 알고리즘은 자연스러운 곡선을 잘 표현하지 못하

고 있다. 이 논문에서 제안하는 영역추출 알고리즘은 4장에서 살펴본 바와 같이 성대의 관심 영역 주위가 관심 영역의 밝기와 비슷한 경우를 제외하고는 대부분의 성대 영상에서 성대의 관심 영역을 잘 추출할 수 있었다. 본 논문에서 제안한 방식을 19명의 성대 환자 영상인 1274개의 프레임에 적용한 결과 95%이상의 성대를 촬영한 영상에서 관심 영역을 추출할 수 있었다.

본 연구에서 제안하는 방법은 각 프레임에서 최소 에너지에 의한 값을 찾아내는 것으로부터 출발한다. 그러나, 어떤 영상은 전반적으로 어두울 수도 있고 이 경우에 관심 영역 이외의 장소에서 최소 에너지가 구해질 수 있다. 남은 연구로는 어떠한 경우에도 특징값을 관심 영역 안에서 효율적으로 찾아내는 방법에 대해 연구하고자 한다.

### 참 고 문 헌

- [1] Rafael C. G., Richard E. W, Digital Image Processing, Addison Wesley, 1993.
- [2] K. V. Mardia, T. J. Hainsworth, "A Spatial Thresholding Method for Image Segmentation," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 10, No. 6, November, 1998.
- [3] 배대규, 한준희, "유전자 알고리즘을 이용한 윤곽선 추출", 한국정보과학회 제24권 제10호, pp. 1063-1072, 10월, 1997.
- [4] 이재현, 장옥배, "움직임 벡터를 사용한 점진적 장면 전환 검출", 한국정보과학회 제3권 제2호, pp. 207-215, 4월, 1997.
- [5] 김진용, 유채권, 황치정, "계층 구조에서 동적 직선 Hough 변환을 이용한 영상 정합", 한국정보과학회, pp. 1073-1081, 10월, 1997.
- [6] 조용진, 양현승, "Snakes 모델을 이용한 얼굴 윤곽선 추적 시스템", 한국정보과학회 제24권 제1호, pp. 62-72, 1월, 1997.
- [7] Laurent D. C., "global Minimum for Active Contour Models: A Minimal Path Approach," International Journal of Computer Vision 24(1), pp. 57-78, 1997.
- [8] A. Hill, T.F. Coates, C.J. Taylor, "Active Shape Models and the shape approximation problem," Image and Vision Computing 14, pp. 601-607, 1996.
- [9] 이재성, 김업준, 박광석, 성미영, 권태영, 모지훈, 성명훈, 김광현, "Videostrobokymography : 성대 운동의 정량적 평가를 위한 새로운 방법 개발", 대한 PACS 학회 제3권, pp. 71-78, 1997.
- [10] L. Brun, J.P. Domenger, "A new split and merge algorithm based on Discrete Map,," Proceedings of the IFIP working group 5.10 on Computer

Graphics and Virtual Worlds - Volume 1, pp. 21-29, 1997.

- [11] Pitas, "Digital Image Processing Algorithm," Prentice Hall.
- [12] R. Depommier, E. Dubois, "Motion Estimation with Detection of Occlusion Areas," Proceedings of the 1992 International Conference on Acoustics, Speech, and Signal Processing, Vol.3, pp. 269-272, 1992.
- [13] Athanasios N., Constantine K. and Ioannis P., "Facial Feature Extraction using Adaptive Hough Transform, Template Matching and Active Contour Models," Proceedings of the 1997 13th Conference on Digital Signal Processing, Vol.2, pp. 865-868, 1997.
- [14] Ronald C., Chi-kin H., "Using 2D Active Contour Models for 3D Reconstruction from Serial Sections," IEEE Proceedings of ICPR '96, pp. 849-853, 1996.
- [15] Amir A. A., Terry E. W. and Ramesh C. J., "Using Dynamic Programming for Solving Variational Problems in Vision," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 12, No. 9, pp. 855-867, 1990.
- [16] Theo P., Yuh-Tay, L., "Integrating Region Growing and Edge Detection," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 12, No. 3, pp. 225-233, 1990.
- [17] Pitas, Digital Image Processing, Prentice Hall, pp.260-274.
- [18] [http://www.ph.tn.tudelft.nl/PRInfo/PRInfo/software/msg\\_00049.html](http://www.ph.tn.tudelft.nl/PRInfo/PRInfo/software/msg_00049.html)



김업준

1997 2월 인천대학교 전자계산학과 공학사. 1999 2월 인천대학교 전자계산학과 석사. 관심분야는 티미디어, 멀티미디어 협동작업, 영상 인식, 음성 인식



성미영

1982 서울대학교 식품영양학과(학사). 계산통계학과(계산학 전공) 부전공. 프랑스 INSA de Lyon 전산학과(공학석사). 1990 프랑스 INSA de Lyon 전산학과(공학박사). 1990 ~ 1993 한국전자통신연구소 컴퓨터연구단 선임연구원. 1993 ~ 현재 인천대학교 전자계산학과 부교수. 관심분야는 멀티미디어 협동작업(multimedia collaborative computing) 멀티미디어 저작, 음성 인터페이스