

초음파 영상을 위한 계층적 특징점 기반 블록 움직임 추출

(Hierarchical Feature Based Block Motion Estimation for
Ultrasound Image Sequences)

김백섭[†] 신성철^{**}
(Baek Sop Kim) (Seong Chul Shin)

요약 연속된 초음파 영상 시퀀스로부터 파노라마 영상을 만들기 위해서는 인접된 프레임 사이의 움직임을 추정해야 한다. 기존에는 고정 블록 움직임 추정 방법이 주로 사용되고 있는데 본 논문은 정확성을 높이고 계산시간을 단축하기 위해 다해상도 영상을 이용한 특징점 기반 블록 움직임 추정 방법을 제안한다. 기존의 블록 움직임 추정 방법은 규칙적으로 블록을 배치하기 때문에 추정된 움직임의 정확도를 높이기 위해서는 블록의 크기가 커지기 때문에 처리 시간이 오래 걸린다. 본 논문에서는 특징점을 중심으로 블록을 배치하여 움직임 추정의 정확도는 유지하면서 블록의 크기를 줄일 수 있었다. 어퍼쳐문제(aperture problem)을 줄이기 위해 코너점을 특징점으로 하였다. 움직임 추정 영역은 일정한 크기의 부영역으로 나누고, 각 부영역에서 가장 코너 강도가 큰 점을 선택하였다. 특징점을 선택하는 데는 해리스 스테판 코너 검출기를 사용하였다.

코너점들이 한 곳으로 편중될 경우 블록들이 움직임 추정 영역에서 끌고루 분산되지 않아 이렇게 구한 블록 움직임을 이용하여 전역 움직임을 구하면 오차가 커진다. 본 논문에서는 이러한 문제를 해결하기 위해 특징점을 선택하는 영역에 제한을 두도록 하였다.

초음파 영상에는 스펙클과 잡음이 많아 코너점을 구하기 전에 영상 평활화를 해야 한다. 계산시간을 줄이고 잡음이 감소된 영상에서 코너점을 구하기 위해 저해상도 영상에서 블록 움직임을 구한 후 점점 고해상도로 확산하는 형태로 다해상도 영상을 사용한다.

실제 세가지 종류의 초음파 영상 시퀀스에 대해 실험결과 제안된 방법은 기존의 방법에 비해 움직임 추정 오차(Displaced Frame Difference)를 평균 66.02에서 58.98로 줄이면서 계산시간은 평균 71ms에서 44ms 으로 빠르게 됨을 알 수 있었다.

키워드 : 초음파 영상, 블록 움직임 추출, 특징점, 코너 검출기, 다해상도 구조, 파노라마 영상

Abstract This paper presents a method for feature based block motion estimation that uses multi-resolution image sequences to obtain the panoramic images in the continuous ultrasound image sequences. In the conventional block motion estimation method, the centers of motion estimation blocks are set at the predetermined and equally spaced locations. This requires the large blocks to include at least one feature, which inevitably requires long estimation time. In this paper, we propose an adaptive method which locates the center of the motion estimation blocks at the feature points. This make it possible to reduce the block size while keeping the motion estimation accuracy. The Harris-Stephen corner detector is used to get the feature points.

The corner points tend to group together, which cause the error in the global motion estimation. In order to distribute the feature points as evenly as possible, the image is firstly divided into regular subregions, and a strongest corner point is selected as a feature in each subregion.

The ultrasound images contain speckle patterns and noise. In order to reduce the noise artifact and reduce the computational time, the proposed method use the multi-resolution image sequences.

· 이 논문은 2005학년도 한림대학교 교비 연구비에 의하여 연구되었음

† 중신회원 : 한림대학교 정보통신공학부 교수

bskim@hallym.ac.kr

** 학생회원 : (주)메디슨 연구소 연구원

neoring@hallym.ac.kr

논문접수 : 2005년 11월 10일

심사완료 : 2006년 2월 21일

The first algorithm estimates the motion in the smoothed low resolution image, and the estimated motion is prolonged to the next higher resolution image. By this way the size of search region can be reduced in the higher resolution image.

Experiments were performed on three types of ultrasound image sequences. These were shown that the proposed method reduces both the computational time (from 77ms to 44ms) and the displaced frame difference (from 66.02 to 58.08).

Key words : ultrasound image, block motion estimation, feature point, corner detector, multi-resolution, panoramic image

1. 서론

의료 영상 시스템은 영상 획득의 목적과 촬영 신체 부위에 따라 자기공명영상(magnetic resonance image), 전산화 단층촬영술(computerized tomography), 초음파 영상 진단기 등 다양한 시스템이 개발되었다. 특히, 초음파 영상 기기는 다른 의료 영상 시스템에 비해, 시스템의 규모가 작고, 신체에 끼치는 영향이 적어서 반복 검사가 가능하고, 실시간으로 단층 영상을 얻을 수 있는 장점을 지니고 있어, 다양한 분야에서 이용되고 있다.

초음파 영상은 촬영하고자 하는 부위의 프로브(probe)로부터 초음파를 내보낸 후, 서로 밀도가 다른 매질에 의해 반사되는 음파를 분석하여 얻어지는 영상이다. 따라서, 촬영된 초음파 영상의 크기는 촬영에 사용된 프로브의 크기에 비례하게 된다. 초음파 파노라마 영상은 근골격계, 갑상선 등의 해부학적 구조와 관계 등을 진단하기에 보다 적합하도록 넓은 시야의 영상(field of view extension)을 조명한 것처럼 표시해주는 방법으로 최근 3, 4년 전부터 초음파 진단기기의 필수 기능으로 채택되고 있다.

파노라마 영상을 얻기 위한 방법에는 응용분야에 따라 다양한 영상처리 기술이 필요하겠지만 크게 영상 정합(registration)과 영상 합성(compounding)기술이 필요하다[1]. 그리고 파노라마 영상이 이루는 면은 영상 취득 모델에 따라 평면, 원통 등의 여러 가지 형태를 가질 수 있으나 대개 평면을 가정한다[2]. 이 경우 한 장의 기준 영상(reference image)을 정하고, 각 영상의 기준 영상으로부터의 상대 위치를 추정하여 배치시킨다.

영상정합을 위해서는 각 연속 영상 간의 움직임을 추정하는 것이 필요하다. 움직임 추정은 정합 영상차(displaced frame difference)를 최소화하는 움직임 모델 요소를 구하는 최적화 문제이다[1,3-5]. 일반적으로 움직임 추정 방법에는 블록 정합법(block matching estimation), 화소 반복법(pel recursive estimation), 광학 흐름(optical flow estimation) 등이 사용된다.

블록 정합법은 영상을 일정한 크기로 겹치지 않게 블록을 분할하고, 현재 블록과 가장 잘 매치가 되는 블록을 이전 영상의 탐색 영역에서 찾는 방법이다. 블록 정

합법은 정확한 움직임을 비교적 적은 계산으로 추정하기 때문에 가장 많이 사용되고 있다. 화소 반복법은 영역 매칭 기술(region-matching technique)의 한 방법으로 블록 정합법과 대조적으로 화소 단위로 각기 다른 움직임 벡터를 반복적으로 추정하는 방법으로 움직임 추정은 이전 이웃 화소를 이용한다. 이 방법은 구현이 복잡하고 계산량이 많다. 광학 흐름 방법은 계산 시간이 오래 걸리고, 변화가 큰 경우 움직임 벡터의 오차가 커진다는 단점이 있지만 다른 방법에 비해 개체의 움직임을 화소단위로 정확하게 추정하기 때문에 최근 비디오 부호화와 물체의 3차원 움직임과 같은 동영상 해석 등 다양한 분야에서 이용되고 있다.

초음파 영상의 경우 초기 파노라마 기술로는 로봇 팔에 프로브를 장착하여 위치 센서를 통하여 하드웨어적으로 움직임을 추정하는 방법(B-arm scanning systems)이 사용되었다. 이 방법은 별도의 하드웨어가 필요하고 위치 센서에서의 전자기적인 에너지 방출이 초음파 에너지에 간섭하게 되는 등 문제가 발생하였다. 이러한 문제를 해결하기 위해 4, 5년 전부터는 프로브의 움직임 추정을 소프트웨어적으로 처리하는 기술이 연구되었다[6-11]. 이런 연구의 한 방법으로 [10]은 프로브에 움직임이 가로 방향 움직임과 회전만 있다고 가정하고, 초음파 영상의 투영 분포(projection distribution)을 통한 1차원 정합(1-D matching)으로 변환하여 움직임을 추정하였다. 이 방법은 빠르게 움직임을 추정하지만, 추정된 움직임에 정확도가 낮은 문제가 있다.

초음파 영상에 블록 정합법을 이용하는 방법[11]은 움직임 추정 영역을 8개의 부영역(sub region)으로 나누고 각 부영역의 움직임을 블록 정합법으로 추정한 후, 각 부영역에서 추정된 움직임을 이용하여 전역 움직임을 추정한다. 여기서 블록의 중심은 현재 프레임에서 각 부영역의 중심으로 고정하고 이전 프레임에서 이 블록과 정합도가 가장 높은 블록의 중심을 구했다.

본 논문에서는 움직임 추정의 정확도를 유지하면서 계산시간을 단축하기 위한 방법을 제안한다. 움직임 추정 영역을 영상에 따라 적응적으로 변경함으로써 블록 크기를 줄일 수 있어 처리 시간을 단축하도록 하였고

움직임 추정 영역에 제한을 두어 전역 움직임 오차를 줄이도록 하였다. 또한 다해상도(multi-resolution)영상을 사용하여 특징점 추출과 블록 움직임 추정을 계층적으로 수행함으로써 처리 시간을 더욱 단축하였다.

블록 정합법으로 움직임을 예측할 때 블록 배치와 블록의 크기가 성능에 중요한 영향을 준다. 블록의 배치는 전역 움직임 요소의 예측 분산에 영향을 주며, 블록의 크기는 계산량에 많은 영향을 준다. 전역 움직임을 예측하기 위하여 움직임 영역 내에 블록을 골고루 배치하는 것이 중요한 반면, 특징이 뚜렷한 영역에서 움직임 추정이 정확하기 때문에 규칙적으로 배치할 경우 특징 영역이 포함되도록 블록을 크게 해야 하므로 계산량이 증가하게 되는 문제가 있다[7]. 이러한 문제를 해결하기 위해 특징점을 중심으로 블록을 배치하여 작은 크기의 블록으로도 특징 영역이 포함되도록 하였다. 특징점은 영상에서 코너(corner)에 배치하여 어퍼추어 문제(aperture problem)을 줄이도록 하였다. 다해상도 영상을 사용하여 먼저 저해상도 영상에서 움직임을 추정하여 보다 고해상도 영상으로 확장(prolongate)하도록 하였다. 저해상도 영상에서는 계산 시간을 줄일 수 있을 뿐 아니라 잡음의 영향을 줄여 특징점 검출에서의 오차를 줄일 수 있기 때문이다.

2장에서 제안된 방법을 설명하고 3장에 실험결과를 제시한 후 4장에서 본 논문의 결론을 내린다.

2. 계층적 특징점 기반 블록 움직임 추정

기존의 블록 정합법은 고정된 부영역을 사용하는데 반해 본 논문에서는 크게 세 가지 방법을 사용하여 움직임 추정 정확도를 향상시키고, 처리 시간을 단축하고자 하였다. 첫째는 블록을 각 부영역에서 특징이 강한 곳에 적응적으로 배치하는 것이다. 이는 블록의 크기를 줄일 수 있어 처리 시간을 단축하면서 정확한 움직임을 추정하고자 하는 것이다. 둘째는 부영역을 서로 간격을 두어 배치하는 것이다. 이는 움직임을 추정하는 블록들이 편중되게 위치하는 것을 방지하여 전역 움직임을 예측하는데 발생하는 오차를 줄이기 위한 것이다. 마지막으로 위의 두 방법에 다해상도 구조를 적용하여 전체적인 처리 시간을 단축하고자 하였다. 전체 처리 과정은 그림 1과 같다.

2.1 특징점 추출

원 영상 (레벨 N-1 영상)에서 특징점을 구하기 위해 영상을 여러개의 부영역으로 나눈다. 각 부구간에 있는 화소들에 대해 코너점을 추출하여 코너의 정도가 가장 큰 화소를 그 부영역의 특징점으로 선택한다. 코너의 정도를 계산하는데는 식 (1)과 같은 Harris와 Stephen의 코너 검출 방법[12]을 사용한다.

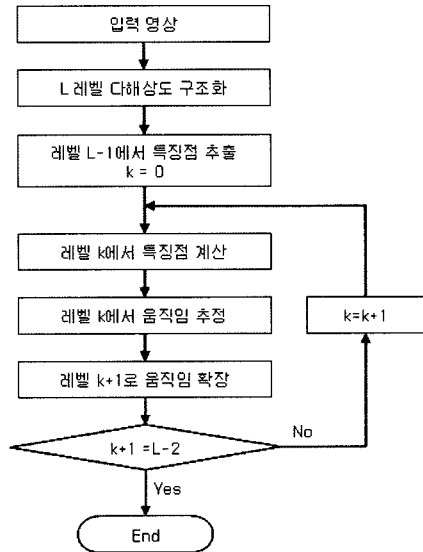


그림 1 전체 처리 과정 흐름도

$$G(i, j) = |N(x, y)| - \kappa Tr^2 N(x, y) \quad (1)$$

여기서, 구조적 행렬(structure matrix) $N(x, y) = \begin{bmatrix} I_x^2 & I_x I_y \\ I_y I_x & I_y^2 \end{bmatrix}$ 이며 Tr 은 행렬의 trace를 나타낸다. I_x, I_y 는 각각 x, y 축에 대한 기울기(gradient)로, 그림 2에 있는 9×9 커널을 사용하여 계산한다. 이 커널은 크기가 3×3 일 때 Prewitt 커널과 동일하다. 초음파 영상은 잡음이 많아 잡음의 영향을 줄이기 위해 크기를 확장하였다. κ 는 0.04로 하였다.

2.2 부영역 설정

정해진 특징점을 중심으로 사각형의 움직임 추정 블록을 설정한다. 움직임 추정 블록이 주어지면 이전 프레임에서 블록을 탐색영역 안에서 움직이면서 현재 프레임과 정합도가 가장 높은 블록을 찾는다. 이렇게 얻어진 두 블록 사이의 움직임을 국부 움직임(local motion)이라고 한다. 초음파 파노라마 영상을 만들기 위해서는 두 프레임 사이의 이동과 회전으로 모델화된 전역 움직임(global motion)을 추정할 필요가 있다. 국부 움직임들로부터 전역 움직임을 계산하는 과정은 다음 절에서 설명한다. 국부 움직임으로부터 전역 움직임을 예측할 때에는 국부 움직임을 계산할 때 사용되는 블록의 배치와 크기가 성능에 중요한 영향을 준다. 움직임 추정 블록은 하나의 부영역에서 한개씩을 설정하는데, 기존의 방법 [11]은 부영역의 중심에 블록을 배치하였다. 이를 고정 위치 방법이라고 하자. 다른 방법으로는 앞 절에서 구한 특징점을 중심으로 블록을 배치하는 것이다. 이를 가변 위치 방법이라고 하자. 그림 3(a), (b)에 각각 고정 위치와 가변 위치 블록 설정 방법에 대한 예를 보인다. 점선

1	1	1	1	0	-1	-1	-1	-1
1	1	1	1	0	-1	-1	-1	-1
1	1	1	1	0	-1	-1	-1	-1
1	1	1	1	0	-1	-1	-1	-1
1	1	1	1	0	-1	-1	-1	-1
1	1	1	1	0	-1	-1	-1	-1
1	1	1	1	0	-1	-1	-1	-1
1	1	1	1	0	-1	-1	-1	-1
1	1	1	1	0	-1	-1	-1	-1

(a)

1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1
0	0	0	0	0	0	0	0	0
-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1

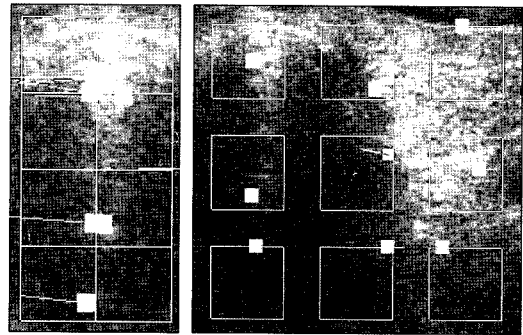
(b)

그림 2 특징점 추출을 위한 9×9 커널. (a) x축 에지 커널, (b) y축 에지 커널

으로 표시한 사각형은 부영역을 나타내고 실선으로 표시한 사각형은 움직임을 추정하는데 사용되는 블록이다. 채워진 원은 블록의 중심을 나타낸다.

국부 움직임 추정 오차를 줄이기 위해서는 블록의 크기는 특징이 뚜렷한 영역을 포함하도록 충분히 커야 한다. 가변 위치 방법은 특징점에 블록의 중심을 두고 있으므로 고정 위치 방법에 비해 더 적은 블록 크기를 사용할 수 있다. 이 결과는 3장 실험 결과에 보인다.

부영역을 그림 3과 같이 서로 인접하도록 설정하면 특징점은 그림 4(a)와 같이 영상에서 구조가 있는 부분에 편중되어 존재하는 경향이 있다. 그림에서 특징점은 채워진 사각형으로 표시하였다. 특징점을 중심으로 국부 움직임을 구한 후 이들 국부 움직임의 정보를 이용하여 전역 움직임을 구하는데, 특징점이 편중되어 있으면 전역 움직임을 추정하는 과정에서 오차가 커진다. 이러한 오차를 줄이기 위해 움직임 추정 영역의 각 부영역 사이에 간격을 두고 배치하여 편중된 블록 배치에 의한 전역 움직임 오차를 줄인다. 이렇게 부영역 사이에 간격을 두고 배치한 것을 간격 부영역 배치라고 하자. 그림 4(b)에 간격 부영역 배치와 각 부영역에서 선택된 특징



(a) 인접 부영역

(b) 간격 부영역

그림 4 부구간 사이 간격 설정

점의 예를 보인다.

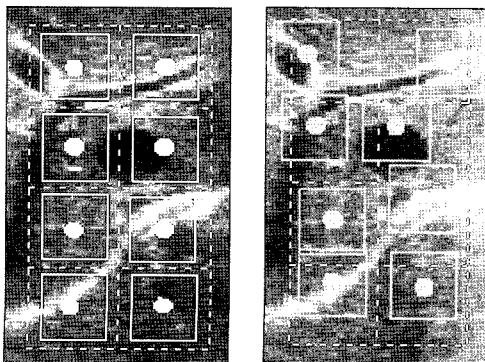
2.3 다해상도 구조에서의 움직임 추정

크기가 $2^N \times 2^M$ 인 원 영상을 $I_{L-1}(N, M)$ 이라고 하자. 여기서 아래 첨자 $L-1$ 은 이 영상의 레벨을 의미한다. 다해상도 구조[14]는 이 영상을 $(I_0(N/2^{L-1}, M/2^{L-1}), I_1(N/2^{L-2}, M/2^{L-2}), \dots, I_{L-1}(N, M))$ 등 레벨 0부터 $L-1$ 까지 L 개의 영상으로 구조화하는 것이다. 레벨 k 의 화소의 명암값 $f_k(i, j)$ 는 레벨 $k+1$ 에서 네 개의 인접한 화소들의 명암값의 평균으로 다음과 같이 bottom-up 방식으로 구한다.

$$f_k(i, j) = \frac{1}{4} \sum_{l=0}^1 \sum_{m=0}^1 f_{k+1}(2i+l, 2j+m), \quad k = L-2, \dots, 0 \quad (2)$$

레벨 k 의 특징점도 원 영상에서 구한 특징점으로부터 bottom-up 방식으로 구한다. 레벨 $k+1$ 의 화소 (i, j) 가 특징점이면 $p_{k+1}(i, j) = true$, 그렇지 않으면 $p_{k+1}(i, j) = false$ 라고 하자. 레벨 k 의 특징점은 다음과 같다.

$$p_k(i, j) = \begin{cases} true, & OR\{p_{k+1}(2i, 2j), p_{k+1}(2i, 2j+1), \\ & p_{k+1}(2i+1, j), p_{k+1}(2i+1, 2j+1)\} \\ false, & otherwise \end{cases}$$



(a) 고정 위치

(b) 가변 위치

그림 3 국부 움직임 추정을 위한 블록 설정 방법

$$, k=L-2, \dots, 0 \quad (3)$$

레벨 k 영상의 한 화소는 레벨 k+1 영상에서 4개의 화소에 대응한다. 또한 한 부영역에는 하나의 특징점만 존재하고, 부영역 사이에 간격을 두었으므로 한 특징점에 인접한 곳에는 다른 특징점이 존재하지 않는다. 그러므로 레벨 k에서의 특징점은 레벨 k+1에서 하나의 특징점에 대응한다.

두 프레임 사이에서 국부 움직임은 현재 프레임에서 하나의 특징점을 중심으로 블록을 배치하고 이전 프레임에서 이 블록과 가장 차이가 적은 블록의 중심점을 탐색을 통해 구한다. 두 블록사이에서의 차이는 블록내 화소들의 차이의 절대값을 합한, SAD(sum of absolute difference)[13]를 사용하였다. 이 때 탐색 영역, 즉 움직임 추정 영역의 크기는 계산량에 많은 영향을 주기 때문에 일반적인 초음파 진단 시 최대의 프로브 움직임을 수용할 수 있는 최소의 범위를 정해야 한다. 일반적으로 초음파 영상에서는 그 특성상 x축 움직임(프로브가 움직이는 방향에 해당)이 많으므로 탐색 범위도 y축 움직임 추정 범위보다 x축 움직임 추정 범위를 크게 설정하였다. 각 축의 탐색 범위는 사용자가 프로브를 움직이는 최대 속도를 실험을 통해 구하여 사용하였다.

움직임 추정은 레벨 0 영상에서부터 레벨 L-1까지 top-down으로 수행된다. 레벨 0 영상에서 움직임을 추정 후 이 움직임을 레벨 1 영상으로 확장(prolongate)하여 레벨 1 에서 움직임을 구할 때 초기값으로 사용하도록 한다. 이 과정을 레벨 L-1까지 반복하여 원 영상에서의 움직임을 구한다. 저해상도 영상에서는 원해상도 영상에서보다 상대적으로 움직임이 적으므로 움직임 탐색 범위를 줄일 수 있어 계산 시간이 적다. 앞에서 설명한 바와 같이 레벨 k에서 한 특징점은 레벨 k+1에서의 한 특징점에 대응한다. 레벨 k에서 특징점의 움직임 벡터를 레벨 k+1에서 대응되는 특징점의 움직임 벡터로 확장된다. 움직임 벡터를 확장할 때는 x, y 축 속도를 각각 2배가 되어야 함을 유의하기 바란다. 이렇게 확장된 움직임 벡터를 레벨 k+1 영상에서 움직임을 구할 때 초기값으로 사용한다. 이렇게 초기값을 설정함으로써 움직임 탐색 영역을 줄일 수 있어 계산 속도를 향상할 수 있다.

2.4 전역 움직임 추정

초음파 프로브를 측면(lateral) 방향으로 움직이고, 두 프레임 사이에 측정 대상물의 움직임을 무시한다면 두 프레임 사이의 움직임은 이동(translation)과 회전(rotation)으로 모델화 할 수 있다. 즉 현재 프레임의 화소의 위치를 $X=(x, y)^t$, 이에 대응하는 이전 프레임에서의 화소의 위치를 $U=(u, v)^t$ 라고 한다면 이들은 다음 식을 만족한다.

$$X=AU+T \quad (4)$$

여기서, $A=\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$ 는 회전 행렬 이며 $T=\begin{pmatrix} t_x \\ t_y \end{pmatrix}$ 이동벡터이다. n개의 국부움직임을 $\{(X_1, U_1), (X_2, U_2), \dots, (X_n, U_n)\}$ 이라고 하면 θ, t_x, t_y 는 다음 식과 같이 제곱 오차를 최소화한 값으로 구할 수 있다.

$$(\theta, t_x, t_y) = \arg \min_{\theta, t_x, t_y} \sum_{i=1}^n |X_i - (AU_i + T)|^2 \quad (5)$$

식 (5)의 해는 다음과 같이 구해진다.

$$\begin{aligned} \theta &= \tan^{-1} \frac{C_{xv} - C_{yu}}{C_{xu} + C_{yv}}, \\ t_x &= \bar{x} - \bar{u} \cos\theta - \bar{v} \sin\theta, \\ t_y &= \bar{y} - \bar{u} \sin\theta - \bar{v} \cos\theta. \end{aligned} \quad (6)$$

여기서 $C_{ab} = \sum_{i=1}^n (a_i - \bar{a})(b_i - \bar{b})$, $\bar{a} = \frac{1}{n} \sum_{i=1}^n a_i$ 로서 각각 공분산과 평균을 의미한다.

3. 실험 및 결과 분석

3.1 실험환경

실제 초음파 영상 시퀀스와 인위적으로 생성한 시퀀스에 대해 각 방법을 비교 실험 하였다. 실험에 사용한 각 방법의 이름을 표 1에 보인다. 표에서 FTR, VTR, VTRG, MVTRG는 각각 Fixed point in Touching Regions, Variable point in Touching Regions, Variable point in Regions having Gap, Multi-resolution Variable point in Regions having Gap을 의미한다. 부영역 설정은 기존의 인접 부영역 방법과 제안된 간격 부영역 방법으로 나누었다. 인접 부영역 방법의 경우 움직임 블록을 설정할 때 부영역의 중심으로 고정된 것(고정위치)과 특징점을 구해 특징점을 중심으로 설정한 것(가변위치)으로 나누었다. 이 경우 단일 해상도 영상을 사용하였다. 기존의 방법은 인접 부영역과 고정위치 블록설정, 즉 FTB방법에 해당한다. 간격 부영역 설정인 경우는 단일 해상도와 다 해상도 영상 구조의 성능 차이에 대해 실험하였다. 이 경우는 가변위치 블록 설정을 사용하였다.

움직임 추정의 정확도를 판단하기 위해서는 정확한 움직임을 알고 있는 검증 데이터(ground truth data)가

표 1 각 방법들의 이름

부영역 설정	블록 설정	영상 해상도	이름
인접	고정위치	단일 해상도	FTR
	가변위치	단일 해상도	VTR
간격	가변위치	단일 해상도	VTRG
	가변위치	다 해상도	MVTRG

필요한데 초음파 영상에서 이를 확보하기는 힘들다. 하나의 방법으로 로봇 팔에 초음파 프로브를 장착하여 기계적으로 정확한 움직임을 가지도록 제어하여 영상과 함께 프로브의 기계적 위치를 저장하여 검증 데이터를 만들 수는 있을 것이다. 이 방법도 팬텀과 같이 강건한 물체에는 사용할 수 있으나 생체는 프로브를 누르는 힘에 따라 영상이 변하므로 현실적으로 생체에 대한 검증 데이터를 만들기는 힘들다. 그러므로 이런 검증 데이터가 없는 환경에서 영상 시퀀스만으로 각 방법의 성능을 평가하는 방법을 사용해야 한다. 본 논문에서는 전역탐색을 통해 얻은 가장 오차가 적은 움직임 벡터를 기준으로 하여 각 방법을 평가하도록 하였다.

초음파 파노라마 영상을 얻을 때는 프로브를 측면 방향으로만 움직이므로 인접한 두 프레임 사이의 움직임은 각도 θ 와 이동량 (t_x, t_y) 등 3개의 요소를 갖는 움직임 벡터를 사용해서 식 (4)와 같이 모델화할 수 있다. X_i, U_i 를 각각 현재 프레임과 이전 프레임에서 i 번째 화소라고 하자. $f_i(X)$ 와 $f_{i-1}(X)$ 는 각각 현재 프레임과 이전 프레임의 명암값을 나타낸다면 두 프레임 사이의 움직임 보상 예측 오차합(sum of displaced frame difference. 이하 SDFD라 함)는 다음과 같다[13].

$$SDFD(\theta, t_x, t_y) = \sum_{i=1}^n |f_i(X_i) - f_{i-1}(AU_i + T)| \quad (7)$$

여기서 $n, A = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$ 와 $T = \begin{pmatrix} t_x \\ t_y \end{pmatrix}$ 는 각각 영상의 전체 화소수, 회전 행렬, 그리고 이동벡터를 의미한다.

본 논문에서는 영상 시퀀스에서 기준 움직임 벡터(reference motion vector)를 추출하여 검증 데이터로 사용하였다. 기준 움직임 벡터의 요소값들은 다음 식으로 구한다.

$$(\theta^*, t_x^*, t_y^*) = \arg \min_{\theta, t_x, t_y} SDFD(\theta, t_x, t_y) \quad (8)$$

실제 계산과정에서는 θ 를 0.1도씩, 그리고 t_x, t_y 를 각각 1화소씩 변경하면서 전역탐색을 통해 SDFD가 가장 작아지는 값을 구하였다. 이는 계산량이 많아 실제 움직임 벡터를 구하는데 사용할 수는 없고 단지 기준 움직임 벡터를 구하기 위한 것이다. 이렇게 구하는 방법을 REF(reference를 의미함)방법이라고 하였으며 표 1의

각 방법을 비교하기 위한 기준으로 사용하였다.

영상 시퀀스는 인체에서 실제 초음파 진단기를 통해 얻은 것 3개(muscle 1개, thyroid 2개)를 사용하였다. 각 시퀀스는 256장의 영상으로 구성되어 있는데 인접된 프레임 사이의 움직임이 너무 적어 한 장씩 건너뛰면서 128장의 영상에 대하여 실험을 수행하였다.

실험 환경은 펜티엄4 1.4Ghz, 메모리 654Mbytes이며 Windows 2000상에서 Visual C++프로그램을 사용하였다.

3.2 실험결과

3.2.1 움직임 추정의 정확도

블록 설정 방법에 따라 움직임 추정의 정확도를 비교하기 위해 FTR 방법과 VTR 방법에 대해 블록의 크기를 각각 13×13, 25×25, 51×51로 변화해 가면서 SDFD를 계산하였다. 그 결과를 표 2에 보인다. 같은 블록 크기에 대해 블록 위치를 가변적으로 설정한 VTR방법이 부영역의 중심으로 고정한 FTR방법 보다 정확도가 높은 것을 알 수 있다. 또한 블록 크기가 작은 경우 두 방법의 차이가 커지는 것을 알 수 있다. 이는 블록 크기가 작은 경우 FTR 방법에서는 블록 안에 구조적인 특징을 가진 화소가 존재할 가능성이 적어져 국부 움직임을 추정하는데 오차가 크기 때문으로 보인다. VTR방법의 경우 블록 크기를 13×13정도로 줄여도 FTR방법에서 25×25를 사용한 경우와 비슷한 결과를 얻을 수 있음을 알 수 있다.

블록 크기를 13×13으로 하여 각 방법의 움직임 추정 정확도를 계산한 것을 표 3에 보인다. 부영역이 서로 인접해 있는 VTR방법보다 부영역 사이에 간격을 둔 FTR방법이 움직임 추정 오차가 적은 것을 알 수 있다. 이는 앞에서 설명한 바와 같이 VTR 방법의 경우 특징점이 몰려 있을 가능성이 높아 전역 움직임을 추정하는데 오차가 커졌기 때문이다.

특이한 것은 MVTRG방법이 VTRG방법보다 움직임 오차가 적다는 점이다. 두 방법은 모두 간격 부영역 설정과 가변위치 블록을 사용하였으나 MVTRG방법이 계산시간을 빠르게 하기 위해 다해상도 구조에서 구현했다는 점만 다르기 때문이다. 이는 초음파 영상의 특성 때문에 나온 결과라고 여겨진다. MVTRG방법에서는 저해상도 영상에서 특징을 구한 후 움직임 벡터를 구해

표 2 블록 크기 변화에 따른 SDFD 차이

데이터	방법	FTR 방법			VTR 방법		
		블록 크기			블록 크기		
		13×13	25×25	51×51	13×13	25×25	51×51
muscle	190.24	78.80	60.13	69.51	62.23	59.77	
thyroid2	112.56	60.72	53.92	70.10	54.40	53.24	
thyroid3	78.27	58.83	55.39	59.24	56.30	55.32	
평균	127.02	66.02	56.48	66.28	57.64	56.11	

표 3 각 방법의 SDFD

방법 데이터	REF	FTR	VTR	VTRG	MVTRG
muscle	55.17	190.24	69.51	65.03	64.66
thyroid2	51.41	112.56	70.10	55.94	51.85
thyroid3	53.45	78.27	59.24	63.84	60.44
평균	53.34	127.02	66.28	61.60	58.98

고해상도로 움직임 벡터를 확장해 나가는 반면 VTRG 방법에서는 고해상도 영상에서 특징점을 구하고 이를 중심으로 움직임 벡터를 구한다. 저해상도 영상은 고해상도 영상을 저주파 필터링 한 것으로 볼 수 있으므로 잡음에 둔감하게 된다. 초음파 영상은 프로브와 전자회로에서 발생하는 잡음 이외에도 특성상 초음파의 간섭에 의해서 스펙클(speckle) 패턴이 존재하게 되어 아주 잡음이 많은 영상이다. MVTRG방법은 잡음 필터링을 수행한 후 특징점을 계산하였기 때문에 보다 구조적인 특징에 해당되는 부분에 특징점을 할당한 것으로 판단된다.

3.2.2 움직임 추정 시간

블록 크기에 따른 계산 시간을 표 4에 보인다. VTR 방법은 FTR방법에 비해 특징점을 추출하기 위한 추가 시간이 필요하다. 표 2에서 설명한 바와 같이 VTR 방법은 13×13 블록을 사용해도 FTR방법 25×25 블록을 사용한 것과 같은 움직임 정확도를 얻을 수 있으나 특징점을 계산하는데 추가적인 시간이 필요하므로 단일 해상도 영상에서는 계산 시간에 이득은 얻을 수 없다는 것을 알 수 있다.

블록의 크기가 13×13인 경우 각 방법의 계산은 표 5와 같다. MVTRG방법은 다해상도 영상을 구성하는데 추가적인 시간이 필요하지만 계산 시간은 전체 시간에 비해 아주 작다. 또한 특징점 추출과 움직임 계산 시간도 VTR나 VTRG방법에 비해 적어져 전체 계산시간이 줄어들음을 알 수 있다. 결국 13×13 MVTRG방법은 25×25 FTR방법에 비해 움직임 추정 오차는 SDFD를 66.02에서 58.98로 줄이면서 계산시간은 71ms에서 44ms으로 빠르게 된다.

3.2.3 파노라마 영상을 통한 주관적 평가

SDFD를 사용해서 추정한 움직임 오차가 적다고 해서 이를 통해 만들어진 파노라마 영상이 반드시 좋다고 할 수는 없으므로 추정된 움직임 벡터를 사용해서 파노

표 5 각 방법의 처리 시간 (단위 : ms)

연산	방법	FTR	VTR	VTRG	MVTRG
다해상도 구조화		0	0	0	2
특징점 추출		0	54	100	22
움직임 계산		22	22	28	20
합 계		22	76	128	44

라마 영상을 만들어 주관적인 평가를 하였다. 움직임을 추정하는데 사용한 블록의 크기는 13×13이다. 그림 5에 muscle 영상 시퀀스에 대한 각 방법의 결과를 보인다. 파노라마 영상의 임상적인 평가는 피부와 혈관과 같은 내부 조직들의 연결성을 보고 판단한다. 13×13 FTR 방법은 움직임 추정 오차가 커서 내부 조직이 연속적이지 않았다. 사각형으로 표시된 부분을 보면 VTR 방법은 하단 부분에 내부 조직이 연속적이지 않은 반면 나머지 방법들은 조직의 연속성이 유지되었다. 원으로 표시된 부분 경우 VTR방법은 REF방법에 비해 영상이 왜곡되는 현상을 보였다. 전반적으로 부 영역 사이에 간격을 둔 VTRG방법과 MVTRG방법이 우수한 결과를 보였다.

4. 결론 및 향후 과제

초음파 영상 시퀀스에서 파노라마 영상을 만드는 기법은 초음파 진단기기의 부가기능 중 하나로 주요 초음파 진단기에 내장되어 있다. 파노라마 영상을 만들기 위해서는 영상 프레임들 사이에 전역적 움직임을 추정해야 한다. 일반적으로 블록 정합법이 사용하여 국부 움직임을 추정한 후 전역적 움직임을 구하는 방법이 사용되고 있다. 본 논문은 국부 움직임을 추정하는데 있어 기존의 방법에 비해 정확도가 높고 속도가 빠른 방법을 제시하였다.

기존 방법은 탐색 블록의 위치를 고정적으로 배치하고 있으나 본 논문에서는 특징점을 기준으로 가변적으로 배치하면 블록 크기를 줄여도 동일한 움직임 추정 오차를 가질 수 있다는 것을 보였다. 특징점이 일부 영역에 편중되어 전역 움직임을 추정하는데 오차가 커질 수 있다. 이 문제를 해결하기 위해 본 논문에서는 서로 간격을 둔 부영역을 설정하고 설정된 부영역 안에서 하나의 특징점을 선택하도록 하는 방법을 제안하였다.

특징점을 중심으로 블록 위치를 가변적으로 설정하면

표 4 블록 크기에 따른 처리 시간 (단위 : ms)

연산	방법	FTR 방법			VTR 방법		
		13×13	25×25	51×51	13×13	25×25	51×51
특징점 추출		0	0	0	54	54	54
움직임 계산		22	71	277	22	71	277
합 계		22	71	277	78	125	331

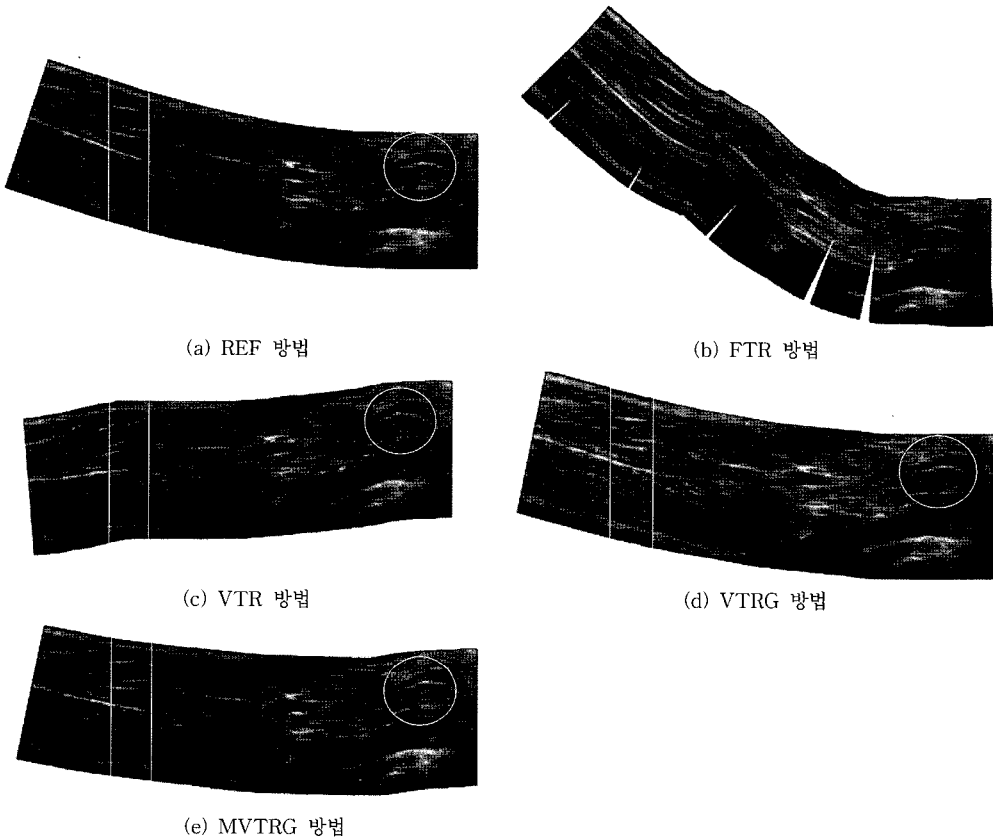


그림 5 각 방법으로 만든 파노라마 영상

탐색 블록의 크기는 줄일 수 있지만 특징점을 계산하는데 드는 부가시간 때문에 전체 계산시간에 대한 이득은 없다. 또한 초음파 영상의 고유한 스펙클 잡음의 영향으로 잘못된 특징점을 선택할 수 있다. 본 논문에서는 이런 단점을 줄이기 위해 다해상도 영상을 사용하는 방법을 제시하였다. 저 해상도 영상에서 특징점을 구한 후 움직임 벡터를 계산해서 고 해상도 영상으로 확장하고 고 해상도 영상에서는 확장된 움직임 벡터를 초기값으로 블록 탐색을 수행하도록 하였다. 이 방법은 잡음이 적은 저 해상도에서 특징점을 구해 오차를 줄일 수 있고 저 해상도 영상에서는 움직임 벡터를 빠르게 구할 수 있으며 이를 초기값으로 고 해상도 영상에서 움직임 벡터를 구할 때는 탐색 영역이 줄어들어 계산시간을 줄일 수 있다.

실제 초음파 영상 시퀀스에 대해 실험결과 제안된 방법은 기존의 방법에 비해 움직임 보상 예측 오차합을 평균 66.02에서 58.98로 줄이면서 계산시간은 평균 71ms에서 44ms 으로 빠르게 됨을 알 수 있었다.

제안된 방법은 움직임 추정에서 hill climbing 방법을 쓰고 있기 때문에 저해상도에서 추정된 움직임 오차를

그대로 고해상도에 전파되는 문제가 있다. 이러한 문제를 해결하기 위해 영상 정합차가 가장 작은 움직임 벡터 하나만 선택하는 대신 여러 개의 후보를 선택하여 고해상도 영상에 전달함으로써 움직임 추정 오차를 더욱 줄일 수 있을 것으로 보인다. 또한 현재는 가장 저해상도 영상에서 특징을 구했으나 영상을 scale space로 변환하여 여러 scale에서 구한 특징들을 혼합하여 사용하면 움직임 추정 오차를 더욱 줄일 수 있을 것으로 판단되어 추후 과제로 남긴다.

참 고 문 헌

- [1] H. Nicolas, "New Methods for Dynamic Mosaicking," *IEEE Tr. on Image Processing*, Vol. 10 (8), pp. 1239-1251, Aug. 2001.
- [2] S. Peleg, B. Rouso, A. Rav-Acha, and A. Zomet, "Mosaicing on Adaptive Manifolds," *IEEE Tr. on Pattern Analysis and Machine Intelligence*, Vol. 22 (10), pp. 1144-1154, Oct. 2000.
- [3] H. Wallin, C. Christopoulos F. Furesjo, "Robust Parametric Motion Estimation for Image Mosaic-

- cing in the MPEG-7 Standard," *ICIP'2001*, Vol. 2, pp. 961-964, Oct. 2001.
- [4] C.-T. Hsu, T.-H. Cheng, R.A. Beuker, and J.-K. Horng, "Feature-based Video Mosaic," *ICIP'2000*, Vol. 2, pp. 887-890, Sept. 2000.
- [5] M. Kourogi, T. Kurata, J. Hoshino, and Y. Muraoka, "Real-time Image Mosaicing from a Video Sequence," *ICIP'99*, Vol. 4, pp. 133-137, Oct. 1999.
- [6] L. Weng, A.P. Tirumalai and L. Nock, "Method and Apparatus for Generating and Displaying Panoramic Ultrasound Images," *U.S. Patent 578 2766*, July 1998.
- [7] T.S. Sumanaweera, L. Pang and S.S. Bolorforosh, "Medical Diagnostic Ultrasound System and Method for Multiple Image Registration," *U.S. Patent 6159152*, Dec. 2000.
- [8] A.L. Robinson, R.R. Entrekin, and J.R. Jago, "Ultrasonic Diagnostic Imaging System with Variable Spatial Compound," *U.S. Patent 6210328 B1*, Apr. 2001.
- [9] J.R. Jago, "Ultrasonic Diagnostic Systems with Spatial Compounded Panoramic Imaging," *U.S. Patent 6416477 B1*, July 2002.
- [10] T. Miyatake, A. Nagasaka and S. Umemura, "Method and Device of Synthesizing Panorama Image from Ultrasound Images," *U.S. Patent 6328693 B1*, Dec. 2001.
- [11] L. Weng, A.P. Tirumalai, and L. Nock, "Method and Apparatus for Generating and Displaying Panoramic Ultrasound Images," *U.S. Patent 5782766*, Jul. 1998.
- [12] C. Harris and M. Stephens, "A Combined Corner and Edge Detector," In *Proc. Alvey Conf.*, pp. 189-192, 1987.
- [13] Yun Q. Shi and Huifang Sun (ed), *Image and Video compression for multimedia engineering*, CRC Press, Ch. 10-13, 1999.
- [14] I. Pitas, *Digital image processing algorithms and applications*, John Wiley & Sons, Ch. 7, 2000.



신 성 철

2003년 한림대학교 컴퓨터공학과 졸업(공학사). 2006년 한림대학교 대학원 컴퓨터공학과 졸업(공학석사). 2006년~현재 (주)메디슨 연구소 근무중(연구원). 관심분야는 영상처리, 컴퓨터비전, 컴퓨터 그래픽스



김 백 섭

1978년 한양대학교 전자공학과 졸업(공학사). 1980년 한국과학기술원 전기 및 전자공학과 졸업(공학석사). 1985년 한국과학기술원 전기 및 전자공학과 졸업(공학박사). 현재 한림대학교 정보통신공학부 교수. 관심분야는 패턴인식, 영상처리,

컴퓨터비전