

論文91-28B-11-4

실시간 이차원 지휘운동의 해석

(On-Line Two-Dimensional Conducting Motion Analysis)

金 鍾 成,* 劉 凡 材,** 吳 尚 錄,*** 卞 增 男*

(Jong Sung Kim, Bum Jae You, Sang Rok Oh, and Zeungnam Bien)

要 約

본 논문에서는 인간의 음악지휘 동작을 시각기능에 의해 실시간으로 관찰하고 이해하는 방법이 제안되었다. 고속의 시각장치를 이용하여 지휘봉의 끝점 위치를 화면상의 좌표로서 추출하고 이 좌표들을 이용하여 지휘도형의 패턴을 인식한다. 음악의 악보를 지휘에 맞추어 연주할 수 있도록 지휘에 대한 전문적 지식에 기초한 알고리즘이 제안되었고, 동시에 센박, 정지상태, 음의 강약의 파악을 위한 알고리즘도 제안하고 있다. 실제 반복실험을 통하여 오차없이 지휘도형을 잘 인식하고 있음을 확인하였다.

Abstract

This paper proposes an on-line method of understanding human's conducting action observed through a vision sensor. The vision system captures images of conducting action and extracts image coordinates of endpoint of the baton. A proposed algorithm based on the expert knowledge about conducting recognizes patterns of the conducting action from the extracted image coordinates and play the corresponding music score. Complementary algorithms are also proposed for identifying the first beat static point and dynamics. Through extensive experiments, this algorithm is found to detect lower edges and upper edges without error.

I. 서 론

시각 센서를 통한 영상의 해석은 크게 정적영상(static image)의 해석과 동적영상(dynamic image)의 해석으로 나눌 수 있다. 동적영상은 정적영상들의

순서적인 집합과 그 순서 및 간격과 관련된 시간함수(time function)로써 정의된다.¹⁾ 정적영상의 해석은 시간에 대한 제약이 별로 없고 처리할 데이터의 양이 한정되어 있는 반면, 동적영상의 해석은 각 정적영상으로부터 추출한 정보와 이러한 영상들 사이의 관계로부터 추출된 정보를 함께 묶어서 해석해야 하므로 정적영상의 경우보다 다루어야 하는 데이터의 양이 방대하여 많은 처리시간이 소모된다.

동적영상의 해석방법은 세가지로 대별할 수 있는바, 이동한 인접 두 영상의 차이(exclusive OR/subtraction)를 이용해서 운동정보를 얻는 차분법(differencing technique), 운동변수(속도, 방향등)를 결정하기 위해 각 영상화소(pixel)에서 시간과 공간에 대한

*正會員, 韓國科學技術院 電氣 및 電子工學科
(Dept. of Electrical Eng., KAIST)

**正會員, 터보테크(株)
(Turbo-Tek Co., Ltd.)

***正會員, 韓國科學技術研究院 制御시스템 研究室
(Control System Lab., KIST)

接受日字: 1991年 6月 26日

기울기 (gradient) 를 계산하는 시공경사법 (temporal-spatial gradient technique), 그리고 인접한 각 영상에서 특징점을 정해놓고 그 특징점을 확인함으로써 운동에 대한 정보를 얻는 매칭 (matching) 방법 등이 있다.¹¹⁾

그러나 움직이는 물체에 대한 실시간 해석 및 인식을 위해 위의 방법들을 사용하기에는 시간적 제약으로 인한 많은 어려움이 있다. 즉, 신뢰할 수 있을 정도의 물체인식 결과를 얻기 위해서는 방대한 양의 계산이 필요하기 때문에 처리 시간이 매우 길고, 운동에 따른 영상잡음 및 신호왜곡 (signal distortion) 등에 강하기 위해서는 보다 복잡한 영상처리 알고리즘이 필요하다.¹²⁾ 이와 같은 어려움은 특별한 목적의 하드웨어를 사용하고 영상 데이터를 간략화시키며 간단하고 안정된 알고리즘을 이용하여 극복될 수 있고, 그에 따라 영상의 온라인 해석이 가능하다.¹³⁾

온라인 해석의 예로써는 필기체 문자 인식, 몸짓 (gesture) 인식, 음성 인식, 악보 인식 등이 있고, 이와 같은 예들의 경우에 인간은 입력되는 영상의 특징들을 빠른 속도로 추출하고 인식할 수 있다.

본 논문에서는 인간의 감성적 표현이 많이 포함된 지휘자의 지휘운동을 대상으로 움직이는 지휘봉의 움직임으로부터 지휘자의 의사를 온라인으로 추출하고 지휘자의 의도에 맞게 음악을 연주하고자 한다. 즉, 곡의 내용(악보)에 관한 정보는 미리 주어지고 지휘자가 지휘봉으로 정해진 지휘도형(박자)을 그려나갈때 지휘운동 해석을 위해 설계된 고속 영상처리 하드웨어와 그를 이용하여 실시간 운동해석 방법을 적용하며, 지휘운동의 특징 (feature) 들을 온라인으로 추출하고 인식할 수 있는 실시간 영상처리 알고리즘 및 인식결과를 음악연주로 나타내는 시스템을 제안하고자 한다.

본 논문의 구성은 제1장의 서론에 이어 제2장에서는 지휘도형에 대한 설명과 온라인으로 지휘도형을 인식하는 알고리즘을 제시하며, 제3장에서는 설명한 알고리즘의 실험결과를 보인다. 마지막으로 제4장에서는 본 연구의 결론 및 앞으로 계속되어야 할 추후 연구과제에 대하여 기술한다.

II. 지휘도형(기법)의 인식

1. 지휘와 지휘도형

지휘는 합창이나 합주에서 지휘자가 음악에 대한 자기의 해석을 지휘봉 또는 맨손을 사용하여 연주자들에게 전달함으로써 연주자들이 지니고 있는 음악에 대한 견해를 통일시켜서 표현하는 행위이다. 최

초의 지휘자의 역할은 곡의 시작을 동시에 하도록 하기 위한 것이었으나, 점차 지휘자의 역할은 다양화되어 곡의 시발을 알려주는 일은 물론 전체 곡의 빠르기 (tempo)와 리듬 (rhythm)을 알려주는 일, 곡의 속도를 이끌어가고 곡의 특정한 부분에서 액센트 (accent)를 지시하는 일 등을 담당하게 되었다.

지휘는 지휘자가 지휘도형을 그리는 동작에 의해 이루어지는 것으로서 지휘도형을 그리는 동작은 시대와 사람에 따라 차이가 있다. (그림1)에서 보는 바와 같이 지휘자는 처음에는 둔탁한 소리를 내서 곡의 시발을 일치시키는 일만 하였고 전체적인 곡의 빠르기를 간단한 원운동으로 표시하기도 하였지만, 지휘도형을 그리는 방법은 계속 변화하여 더욱 섬세하고 정확한 기술로 발전되어 지휘자가 지휘도형을 그리는 기본적인 원리에 따라 지휘를 하면서도 다양한 몸짓을 취하게 되었다.¹⁴⁾

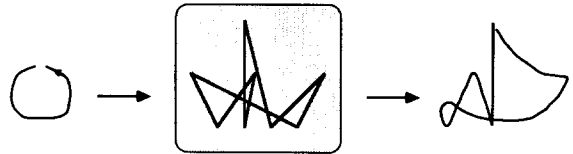


그림 1. 지휘도형의 변화

Fig. 1. Transition of conducting diagram.

지휘도형에 관련된 몇가지 용어를 알아보면 다음과 같다. 첫째, 박 (beat)을 구분하는 정점 (upper corner), 박점 (lower corner)이 있고, 둘째, 각 마디의 첫번째를 표시하는 지휘도형의 중앙 위에서 아래로 강하게 (빨리) 움직이는 선평의 운동이 있다. 셋째, 곡의 셈과 여림을 표시하기 위하여 지휘도형 전체의 크기를 변화시킬 수 있는데 곡의 중간에 $f(p)$ 와 같은 표시가 나타날 때는 지휘도형의 크기를 크게 (작게)하고 그에 비례하여 속도를 빨리 (느리게) 함으로써 전체 빠르기에는 변화가 없게 한다. 곡의 중간에 늘임표가 있을 때는 지휘봉을 정지함으로써 그 음을 2-3배 늘여서 연주할 수 있다.¹⁵⁾

일반적으로 지휘자는 양손을 사용하여 다양한 음악을 표현할 수 있지만 본 논문에서는 다음의 가정 하에서 지휘봉 운동의 온라인 해석을 수행한다.

- 1) 지휘자의 지휘운동은 한손으로 이루어진다.
- 2) 지휘도형의 모양은 지휘원리에 입각한 직선화된 기본 지휘도형으로 한다(그림1의 빗금친 부분)

3) 지휘도형은 카메라 인식 영역내에 전체가 존재한다.

4) 지휘봉의 운동을 쉽게 구분할 수 있도록 지휘봉의 끝에 배경과 구분되는 물체를 단다.

이상과 같은 실험상황에서 움직이는 지휘봉의 동작은 고속의 시각 시스템으로부터 들어오는 지휘봉의 중심좌표에 의해 실시간(real time)으로 추적된다. 앞에서 설명한 지휘도형의 원리를 기초로 하여 지휘도형을 해석하기 위해 다음과 같은 특징들(features)을 추출한다.

- 첫째, 지휘도형에 반드시 나타나는 정점과 박점,
- 둘째, 곡 전체의 리듬(rhythm)을 잃지 않고 따라가기 위한 첫번째 고동박의 예비박(각 마디의 첫부분)
- 셋째, 곡의 중간에 늘임표가 있거나 운동 속도가 매우 느린 경우를 위한 정지된 상태,
- 네째, 곡의 중간에 셈, 여림등의 표시(mf, f, mp, p...)가 있을 때 지휘자에 의한 전체 지휘도형의 크기변화

그와 함께 안정된 영상처리를 위하여 영상잡음에 의한 영향을 최소화 한다. 언급된 특징치 추출은 먼저 지휘도형의 특성상 운동의 방향변화가 생기는 부분을 찾음으로써 가능하고, 미리 정해둔 잡음의 형태를 들어오는 데이터의 변화와 비교함으로써 온라인으로 지휘도형상의 잡음을 제거할 수 있다. 또한 센 박은 지휘도형에서 가장 높은 부분에 있는 정점을 찾으면 되고 정지상태는 들어오는 데이터가 변화하지 않음을 보고 확인할 수 있다. 위의 방법에 대하여 상세히 설명하면 다음과 같다.

2. 정점, 박점의 인식

지휘자가 지휘도형을 그려나갈 때 가장 중요한 요소 중의 하나가 정점과 박점의 명확성이므로 이들의 파악은 매우 중요한 의미를 가진다.

(그림2)에 있는 지휘도형에 나타난 여러가지 형태의 정점과 박점 부분을 보면 알 수 있듯이 이 부분을 찾기 위해 기술기의 급격한 변화를 이용할 수도 있지만 그 과정 중에 삼각함수의 계산 등이 들어가기 때문에 좀 더 빠른 동작을 온라인으로 분석하기 위해서 계산이 간단한 새로운 특징치(feature)로서 방향변화(direction change)를 이용한다.

즉, 실시간으로 들어오는 중심의 x, y좌표 중 y축 데이터를 비교하여 방향 변화가 생긴 부분을 정점 또는 박점이라고 파악한다. 이러한 방향 변화를 확인하기 위해 y축의 최근 세점만을 이용하는 경우에는 지휘중에 (그림3)과 같은 잡음(noise)이 발생할 경우 실제의 정점, 박점을 찾을 수 없게 된다.

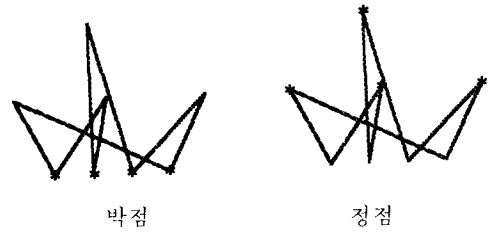


그림 2. 지휘도형의 정점과 박점

Fig. 2. Upper corner and lower corner of conducting diagram.

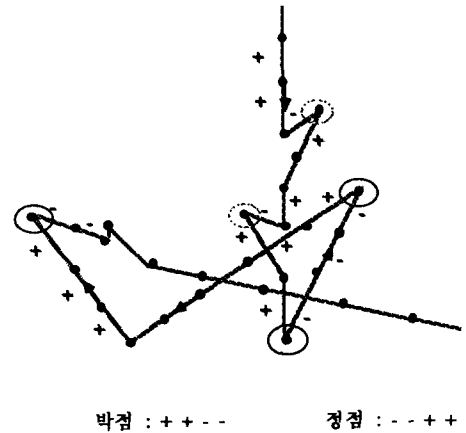


그림 3. 정점과 박점의 인식

Fig. 3. Recognition of upper and lower corner.

따라서 이러한 잡음에 둔감하면서 그리고 많은 시간지연(delay) 없이 정점과 박점을 찾기위한 방법으로 최근에 얻어진 다섯점의 y축 좌표를 이용하고자 한다. 즉, 실시간으로 들어오는 데이터 중 최근 다섯개를 가지고 그 중 y축 데이터의 변화(direction change)가 ++--이면 박점이 생긴 것이라고 보고 --++이면 정점이 생긴 것이라고 본다. 이와 같은 경우 정점과 박점을 인식하는 시점이 항상 두개의 샘플링이 지난 순간이므로 정점과 박점을 찾는 순간은 실제와 비교하면 약120[msec] 정도의 시간지연이 생기지만 전체적인 지휘운동의 온라인 분석작업에 큰 영향을 미치지 못하는 것이다. (그림3)은 위의 방법을 도식적으로 설명한 것이다.

3. 잡음(noise)의 제거

앞 절에서 설명한 것처럼 최근 다섯점을 이용할 경우에 정점, 박점이 아닌 부분에서는 잡음에 둔감하지만 (그림4)에서 나타난 바와 같이 정점, 박점 부분

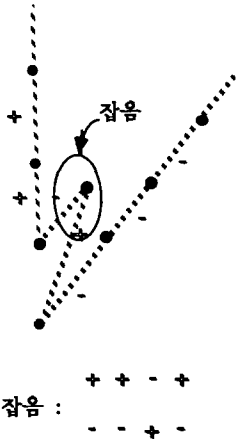


그림 4. 잡음의 제거
Fig. 4. Reduction of noise.

에서 잡음이 발생했을 경우는 실제의 정점, 박점을 찾지 못하게 된다.

따라서 실시간으로 들어오는 최근 다섯점을 미리 정해진 형태 (+ + - + 또는 - - + -)의 부호변화가 생기면 잡음이 들어온 것이라고 보고 그 부분을 제거함으로써 모든 경우에 원하는 정점과 박점을 찾을 수 있게 된다. 즉, y축 방향 변화가 + + - +인 경우는 (minus)를 잡음으로 보아 이 점을 제거하고, 마찬가지로 - - + -인 경우는 + (plus)를 잡음으로 보고 제거한다. (그림4)는 위의 방법을 도식적으로 설명한 것이다.

4. 정지상태의 확인

지휘자가 지휘를 하다가 곡의 중간에 늘임표(fermata)가 있으면 지휘봉을 정지한 상태로 있게 된다. 또한 매우 느린곡을 지휘할 때 지휘봉의 움직임이 느려져서 고속 시각시스템(sampling time; 1/15[sec])의 동작 시간보다 느리게 움직이면 같은 영상을 반복적으로 얻게 되므로 정지상태와 같이 인식된다. 온라인 수행을 위해 최근의 다섯점을 이용하고 있으므로 이 다섯점이 지휘도형의 궤적에 대하여 최대한의 정보를 가지고 있도록 하여야 한다. 따라서 정지상태동안의 데이터는 하나만 갖고 있도록 하고 그 후의 변화된 데이터가 나올 때까지의 데이터를 무시하도록 한다.

5. 셀박(각 마디의 첫번째)의 확인

곡이 진행되는 도중에 갑자기 리듬을 잃어 버렸을 때 늦어도 다음번 마디의 첫번째에서는 다시 정상상

태로 돌아오기 위해서 지휘도형의 첫번째 박의 예비박을 찾아야 한다. (그림5)의 지휘도형에서 보는 바와 같이 정점중에서 가장 높은 부분에 있는 정점이 곡의 각 마디의 첫번째 음을 지정하는 예비박이 된다. 따라서 지휘도형의 궤적에 대한 데이터가 실시간으로 주어질때 앞에서 언급한 정점을 찾는 방법으로 정점을 확인한 후에 정점의 y축 데이터가 미리 실험에 의해 정해진 값보다 작으면 이 정점이 찾고자하는 예비박임을 알 수 있다.

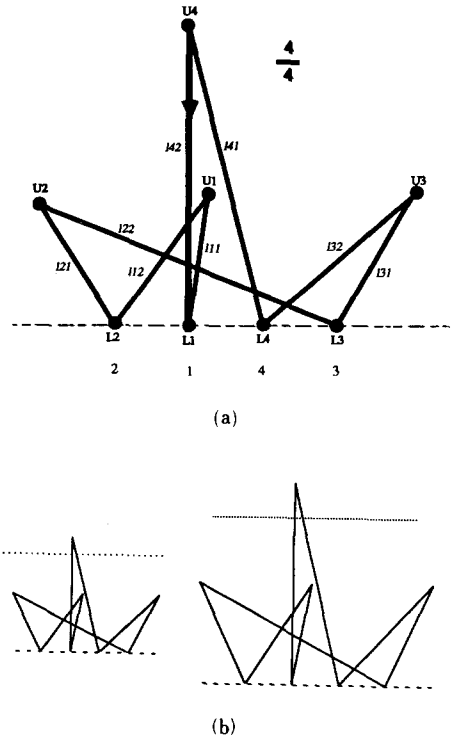


그림 5. (a) 지휘도형의 링크
(b) 지휘도형의 크기변화
Fig. 5. (a) links of conducting diagram,
(b) size change of conducting diagram.

6. 곡의 셈, 여림의 결정

지휘자가 지휘운동을 하다가 곡의 중간에서 mf, f, mp, p 등의 셈, 여림을 표시하는 기호를 만나면 지휘도형의 전체적인 크기를 조절하여 변화를 표현한다. 따라서 인접한 박점사이(link)의 길이의 기준값(reference)을 지휘영역에 근거하여 정한 후에 각 링크(link)의 장, 단에 따라서 곡의 셈, 여림을 나타낼 수 있다. 그림 5(a)는 각 링크의 기준과 비교하기 위해 지휘도형의 각 링크에 이름을 붙인 것이고 그림 5(b)

는 지휘도형의 크기변화를 나타낸 것인데 왼쪽의 작은 지휘도형이 오른쪽보다 더 어린 박을 표시한다.

III. 실험 및 검토

1. 사용된 시각장치의 구성

본 논문에서 사용한 시각장치의 구성은 (그림6)과 같다. 카메라(camera)로부터 들어온 아날로그 영상은 영상 디지털라이저(image digitizer)에 의해서 디지털 영상으로 바뀌고 영상 연산자(LUT/ALU)에 의해서 이진영상(binary image)이 되며, 이진영상 프로젝션 프로세서(BIP)는 2차원 평면상에 있는 이진영상을 고속으로 세개의 축(x축, y축, z축)에 대해 투영(projection)한 결과를 얻고 이것을 이용하여 지휘봉의 중심을 실시간으로 인식한다¹⁾ 이 데이터는 직렬통신(RS-232C)으로 IBM-PC에 전달하게 된다. 시스템의 주 프로세서로는 모토롤라(Motorola)의 16비트 마이크로프로세서인 MC68000 CPU를 사용하였고, MC68881 보조연산 프로세서(coprocessor)를 내장하였다. 시스템의 클럭(clock)은 10MHz이다.

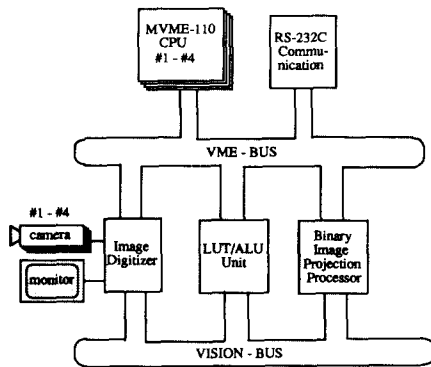


그림 6. 시각장치의 구성
Fig. 6. Construction of vision system.

2. 실험결과 및 검토

실험은 4/4 박자에 대한 직선화된 지휘도형을 대상으로 하였다. (그림7)의 전체시스템에 보인 것처럼 25mm의 촛점 거리를 가지는 CCD 카메라로부터 약 2m 떨어진 곳에 70cm×70cm의 지휘영역을 설정하고 그 범위 안에서 정해진 지휘동작을 행하도록 했다. 앞에서 가정으로 설정했듯이 이진 영상을 사용하므로 지휘봉은 잡음(주위의 불빛)의 영향을 최소화 하기위해 주위가 어두운 상황에서 손전등을 지휘봉으로 사용하여 시각 시스템이 좀 더 명확히 지

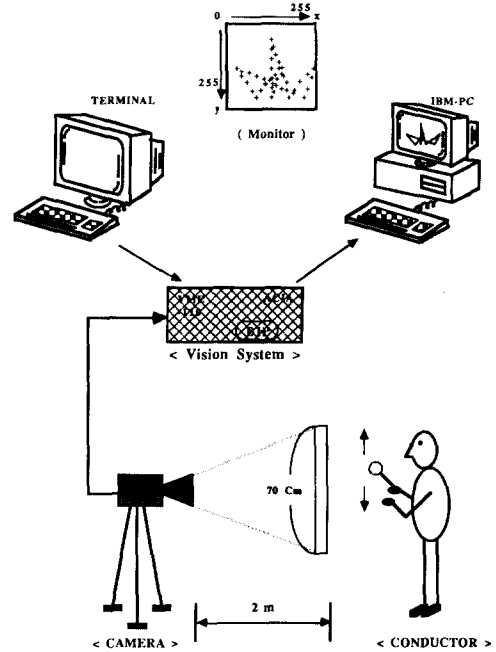


그림 7. 전체 시스템의 구성
Fig. 7. The configuration of total system.

휘봉의 궤적을 추적할 수 있도록 하였다.

처음에 지휘자가 정해진 지휘영역내에서 지휘도형을 그려 나가면 지휘봉의 중심좌표는 고속의 시각장치를 사용하여 구해지고 그 값을 RS-232C 직렬통신(serial communication)에 의해 IBM-PC로 보낸다. 이때의 전송 속도가 9600baud rate(bits/sec)이고 실제로 보내는 데이터는 40비트(bits) 정도이므로 전송할때 생기는 시간지연은 거의 무시할 수 있다. 지휘자가 지휘를 시작할 때 함께 동작을 시작한 IBM-PC는 계속해서 지휘봉의 위치 데이터를 받으면서 모니터(monitor)에 온라인으로 지휘운동의 궤적을 도시하고 앞에서 설명한 알고리즘에 따라 일을 수행한다. (그림8)은 위의 과정을 블록선도(block diagram)로 표시한 것이다.

한 화면의 영상에 대하여 각 단계에서 걸린 시간을 살펴보면 다음과 같다.

Projection 시간 : 30[msec]

Area-center 계산 및 알고리즘 수행 시간 : 30 [msec]

총 수행시간 : 60[msec]

전체 시스템을 수행할 때 다섯점을 이용하므로 정점과 박점의 인식이 두화면 이후에 확인되고 따라서

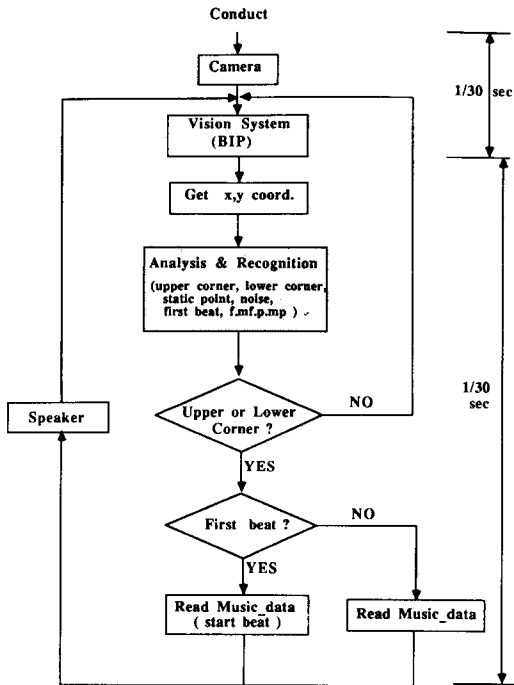


그림 8. 전체 시스템의 블록선도
Fig. 8. System block diagram.

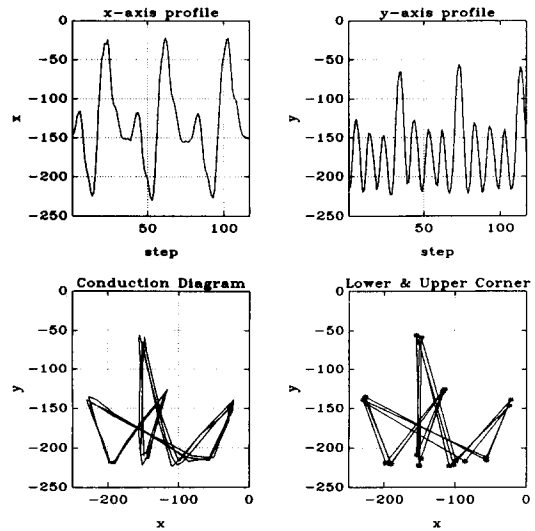


그림 9. 실제의 지휘도형
Fig. 9. Conducting diagram.

약 120[msec]의 시간 지연이 생기지만 온라인동작에 영향을 미치지 않는 짧은 시간이다. (그림9)는 곡의 세마디동안 지휘봉의 실제운동에 대한 x축, y축에 대한 profile과 그 지휘운동의 궤적 및 정점과 박점을 인식한 결과를 보인 것이다.

악보에 대한 정보는 주어져 있다고 가정했기 때문에 4분의 4박자 곡인 ‘학교 중’과 ‘태극기’를 이용하여 얻어진 실험결과를 설명하면 다음과 같다.

먼저 온라인으로 들어오는 데이터 중에서 최근의 다섯점을 이용하여 특징들을 얻었으므로 이 방식에 대한 신뢰도를 직접 실험을 통하여 살펴보았다. (표 1)은 이것의 실험 데이터를 보인 것이다.

여기서 SCH70, SCH90, SCH120은 ‘학교중’에 대하여 M. M.에 의한 빠르기 순서로 지휘한 것이고, FLAG1은 ‘태극기’에 대하여 SCH1과 같이 보통의 빠르기로 지휘를 했을 때 각 방식(P3, P5, P7, P9: Pi는 i개의 최근점을 이용하는 방식을 표시함)에 따른 정점과 박점의 인식 결과를 실제와 비교한 것이다. 여기서 M. M.은 곡의 빠르기표로서 1분 동안 연주할 수 있는 4분음표의 수를 나타낸다. 위 실험 결과에서 보는 바와 같이 최근 다섯점을 이용한 방식이 실제와 가장 잘 부합함을 알 수 있다.

가지고 있는 데이터가 많으면 잡음에는 둔감해지지만 빠른 곡을 연주할 경우는 정점과 박점 사이에 존재하는 데이터의 수가 줄어들어서 실제의 정점, 박점을 놓치게 된다. (그림10)은 위와 같은 경우의 실제 예를 보인것인데 그림의 왼쪽에 있는 궤적의 세번째 고동박에는 박점과 정점 사이에 한개의 점만이 존재하므로 (표1)의 P7 또는 P9의 방법으로는 정, 박점을 제대로 인식하지 못하게 된다.

또한 가지고 있는 데이터가 적으면 빠른 곡에 대해서도 정확한 정점과 박점을 찾을 수 있지만 잡음에 매우 민감해져서 특히 정지상태가 존재하거나 매

표 1. 정점과 박점의 인식

Table 1. Recognition of upper and lower corner.

Data	P3	P5	P7	P7	실제
SCH1	69	66	64	64	66
SCH70	66	66	65	65	66
SCH90	66	66	65	52	66
SCH120	69	66	62	27	66
FLAG1	66	66	58	41	66

(단위 : 정점과 박점의 total 개수)

- Where P3: 최근 3 points로 check
- P5: 최근 5 points로 check
- P7: 최근 7 points 7 check
- P9: 최근 9 points로 check

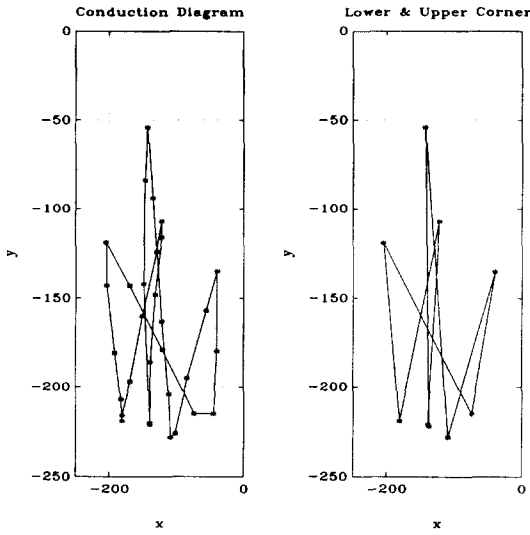


그림 10. 실제의 지휘도형(박점과 정점사이에 데이터의 수가 작은 경우)
 Fig. 10. Conducting diagram(few data).

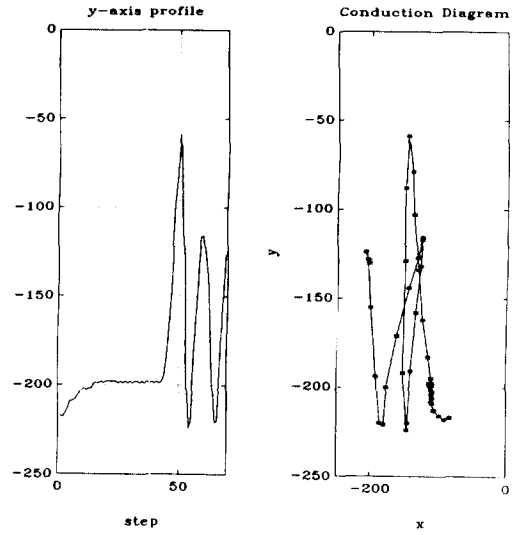


그림 11. 실제의 지휘도형(박점과 정점사이에 데이터의 수가 많은 경우)
 Fig. 11. Conducting diagram(many data).

우 느린 곡을 지휘할 때 물체주위에서 일어나는 aliasing에 의해 1~2화소 정도의 오차가 생기고 이것에 의한 방향 변화에 대해서도 정점, 박점으로 잘못 인식하는 경우가 생긴다. (그림11)은 위와 같은 경우의 실제 예를 보인 것이다. 그림의 왼쪽에 있는 y 축 profile에서 나타낸 바와 같이 초기의 정지 상태에서 출발할 때 많은 횟수의 y축에 대한 방향 변화가 생겼음을 알 수 있다. 즉 P3의 방법은 지휘자의 정지 상태에서의 떨림, 또는 시스템에서 생기는 미세한 변화 등 많은 잡음의 요소에 민감함을 알 수 있다.

곡의 빠르기는 빠르기 표(M.M.)에 의해 나타낼 수 있으므로 같은 곡('학교중')에 대하여 빠르기를 다르게 하면서 반복실험함으로써 성능의 열화없이 지휘와 연주를 할 수 있는 한계를 조사하였다. (그림12)는 빠르기를 다르게 하면서 같은 측정 시간(30 steps)동안 지휘운동의 궤적을 그린 것이다. 여기서 괄호안에 적힌 숫자(50, 80, 110, 130)는 M.M.으로 표시한 빠르기이다. (그림13)에서 알 수 있듯이 곡이 빨라질 수록 정점과 박점 사이에 존재하는 점의 수가 적어짐을 알 수 있다. 위 실험의 결과 및 빠르기 표에 의한 시간계산을 통해 현재의 시스템으로 연주할 수 있는 최대의 빠르기는 M.M. ♩=140~150 정도임을 알 수 있었다. 참고로 이 빠르기로 '학교중'을 지휘할 경우 전체곡을 연주하는데 약 12~14[sec]

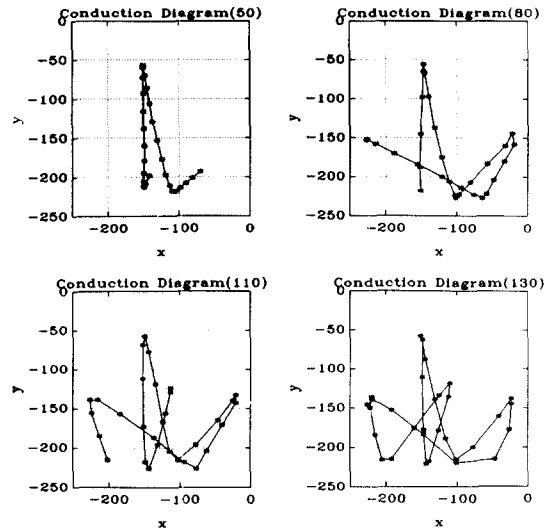


그림 12. 빠르기의 변화에 따른 지휘도형
 Fig. 12. Conducting diagram about velocity change.

가 걸린다.

이제 음악 데이터의 형식(format)에 대하여 설명하고자 한다. 곡에 대한 정보가 주어져 있다고 가정했기 때문에 IBM-PC에 곡의 정보를 주고 적절한 소리를 내기 위해서 각 음표에 대해서 3가지의 내용

384	2	1
384	2	0
427	2	0
427	2	0
384	2	1
384	2	0
320	4	0
384	2	1

⋮
⋮
⋮
⋮
⋮

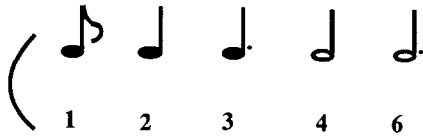


그림 13. 음악 데이터와 박의 길이
Fig. 13. Music data and length of beat.

을 가지도록 했다. 첫번째는 피치(pitch)로서 각 음의 계명에 해당하는 주파수를 나타내고, 두번째는 8분음표를 기준으로 정해서 박의 길이를 표시하며, 세번째는 해당 음표가 각 마디의 첫번째 박(센박)인가를 알려주도록 0(여린박)/1(센박)로 나타내었다. (그림13)은 ‘학교 중’에 대한 음악 데이터의 예와 8분음표를 기준으로 한 박의 길이를 나타낸 것이다.

위와 같은 음악 데이터가 있을 때 지휘자가 지휘를 시작하면 실시간으로 들어오는 데이터를 분석한다. ‘학교 중’의 경우를 예로 들어 설명하면 곡의 시발을 표시하는 박점이 나타나면 그때의 음악 데이터가 ‘384 2 1’이므로 첫번째 요소인 음계 ‘솔(384 Hz)’을 소리내면서 동시에 두번째 요소가 ‘2’를 표시하므로 한개의 고동박 즉, 다음 박점이 나올때까지 앞에서 언급한 알고리즘을 수행하고 다시 새로운 박점이 나오면 음악 데이터를 읽고 반복된 일을 수행한다. 여기서 만약 음악 데이터의 두번째 요소가 ‘1’이라면 곧바로 다음번 정점에서 다음번의 음악 데이터를 읽게 된다. 만일 곡의 중간에서 한박을 놓치게 되더라도 새로운 마디가 시작되는 첫번째 박에서는 각 마디의 첫번째 박을 인식하고 음악 데이터의 세번째에 1이 있을 때까지 해당 데이터를 소리내지 않고 넘어간다. 따라서 곡의 리듬을 잠시 놓쳤을 때에도 즉시 원래의 리듬을 찾을 수 있다.

본 논문에서는 지휘자의 지휘운동에 따라 소리를 내는 매개체로서 IBM-PC의 스피커(speaker) 진동을 이용하기 때문에 해당 음표의 계명과 길이만을 나타낼 수 있고 곡의 셈,여림을 직접 출력으로 들을 수는 없지만 요즈음 많이 사용되고 있는 MIDI(musical instrument digital interface)나 PC music card와 연결하면 좀 더 완벽한 음악을 들을 수 있을 것이다.

IV. 결 론

본 논문에서는 인간의 여러가지 행동 중에서 특히 감성적 부분이 많이 포함된 지휘자의 지휘운동을 공학적인 측면에서 주요한 특징들을 고속의 실시간 영상처리 시스템을 이용하여 추출하고 이 정보들로부터 온라인으로 음악을 연주할 수 있음을 보였다. 즉, 공간에서 정해진 모양의 지휘도형을 그리면서 지휘를 할때 고속의 시각장치에서 그 지휘운동의 궤적을 실시간으로 추적하면서 지휘운동이 표현하고자 하는 바를 방향 변화(direction change)를 특징으로 추출하여 정형화된 음악 데이터와 함께 온라인으로 음악을 연주할 수 있게 하였다. 본 논문에서는 지휘도형의 원리에 근거하여 정해진 지휘영역내에서 지휘운동을 수행할 때 주위의 잡음에 둔감하도록 온라인으로 잡음 성분을 제거하였다. 또한 지휘자가 속도(빠르기)를 달리하면서 지휘를 할 때에도 하드웨어의 한계(M. M. J = 150)까지는 잘 따라가면서 음악을 연주할 수 있음을 확인하였다. 이 빠르기는 ‘보통 빠르게(Moderato)’보다 약간 빠른 것이므로 대부분의 곡에 적용 가능하리라고 본다. 그리고 본 연구에서는 지휘가 진행되는 동안 곡의 중간에서 리듬을 놓쳤을 때 즉시 원래의 리듬을 찾을 수 있게 하였고 곡의 셈,여림에 대하여 지휘도형의 크기 변화를 확인함으로써(실제로 음악으로 표현은 할 수 없었지만) 그것에 따라 악곡의 진행중에 나타나는 곡의 셈,여림을 알아낼 수 있었다.

그러나 전체적인 알고리즘이 다섯점을 이용하기 때문에 원칙적으로 정점과 박점을 포함하여 그 사이에 세점 이상이 존재하지 않으면 제대로 곡을 연주할 수 없다는 것이 현재의 시스템과 알고리즘의 한계이다.

그리고 구현된 시스템은 온라인으로 들어오는 최근의 다섯개의 데이터를 중심으로 운동을 해석하기 때문에 약 120[msec] 정도의 시간지연이 존재하지만 정, 박점의 위치가 어느 범위안에 존재하므로 예측 방법을 이용하여 지금보다 더 빠른 응답을 얻을 수 있을 것이다. 그리고 곡의 셈, 여림을 표시하기

위해서 미리 정해진 기준값을 주고 그 값의 크기에 따라 셈, 여림을 결정하였지만 초기의 기준값 설정 없이 몇번의 반복된 실행을 통해서 이를 결정할 수 있을 것이다.

본 논문의 결과를 이용하여 여러 종류의 악기에 대한 악보가 주어졌을 때 그 악보들을 시각장치를 통해서 인식하고 그 데이터를 음원에서 사용할 수 있는 형식으로 바꾸며, 일반적인 지휘도형 (곡선화면도형)에 대해 - 결국 일정시간(1/15 sec) 간격으로 샘플링(sampling)되므로 직선화면 지휘도형과 같다. 온라인으로 지휘자의 지휘에 따라 음원(Music Card, MIDI 등)을 통해 음악을 연주하는 시스템을 구현할 수 있을 것이다.

參 考 文 獻

- [1] W.N. Martin and J.K. Aggarwal, 'Survey Dynamic Scene Analysis,' *Computer Graphics and Image Processing*, vol. 7, pp. 356-374, 1978.
- [2] T. Uno, M. Ejiri and T. Tokunaga, "A Method of Real-Time Recognition of Moving Objects and its Application," *Pattern Recognition Pergamon Press*, vol. 8, pp. 201-208 1976.
- [3] 최승식, "확장 영상 공간에서 정의된 궤적에 의한 로봇의 추적 시스템," 한국 과학기술원, 석사 논문, 1990.
- [4] B.K. Paul Horn, *Robot Vision*, MIT Press, 1988.
- [5] A. Rosenfeld, A.C. Kak, *Digital Picture Processing Academic Press*, 1982.
- [6] D.H. Ballard and C.M. Brown, *Computer Vision*, Prentice-Hall
- [7] N. C. Mohanty, "Computer Tracking of Moving Point Targets in Space" *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-3, no. 5, Sep. 1981.
- [8] J.W. Roach and J.K. Aggarwal "Determining the Movement of Objects from a Sequence of Images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-2, no. 6, Nov. 1980.
- [9] M.D Levine Y.M. Youssef and A. Boyarsky, 'The Quantification of Blood Cell Motion by a Method of Automatic Digital Picture Processing' *IEEE Trans Pattern Anal. Mach. Intell.*, vol. PAMI-2, no. 5, Sep. 1980.
- [10] J.K. Aggarwal and R.O. Duda "Computer Analysis of Moving Polygonal Images" *IEEE Trans Computer* vol C-24, pp. 966-976 Oct. 1975.
- [11] Joonki Kim, 'On-line Gesture Recognition by Feature Analysis' *Computer Vision and Shape Recognition World Scientific*, 1989.
- [12] S.Ando and T. Goto, "Current Status and Future of Intelligent Industrial Robots" *IEEE Trans on Industrial Electronics*, vol. IE-30, no. 3, Aug. 1983.
- [13] Michael Tischer, *PC System Programming for Developers*, Data Becker Book, pp. 447-456, 1989.
- [14] 김원주, "인쇄악보 인식 시스템의 구현에 관한 연구," 한국 과학기술원, 석사학위 논문, 1987.
- [15] M.M. Miller *Introduction to Muisis*, Barnes & Noble Books 1958.
- [16] Musical Theory, 서울, 아름 출판사, 1987.
- [17] 박봉석, 최신 음악 통론, 서울, 아가페 음악 선교원, 1986.
- [18] 이택희, 지휘 도형의 이해, 서울 기독교 음악사 1986.
- [19] 이택희, 합창 지휘법 연구(1)(2), 서울, 한길 예술원, 1987.
- [20] 김종성, "실시간 이차원 운동의 해석," 한국 과학기술원, 석사학위 논문, 1991.

 著 者 紹 介



金 鍾 成 (正會員)

1966年 10月 17日生. 1989年 2月
 경북대학교 전자공학과 학사학위
 취득. 1991年 2月 한국과학기술
 원 전기 및 전자공학과 공학 석
 사학위 취득. 1991年 3月~현재
 한국과학기술원 전기 및 전자공
 학과 박사과정 재학중. 주관심분야는 패턴인식, 보로
 트 비전, 영상처리 등임.

吳 尚 錄 (正會員) 第27卷 第6號 參照

현재 한국과학기술연구원 제
 어시스템연구실 선임연구원

◆

卞 增 男 (正會員) 第28卷 第7號 B編 參照

현재 한국과학기술원 전기
 및 전자공학과 교수

◆

劉 凡 材 (正會員) 第28卷 第7號 B編 參照

현재 (주)터보테크 기술연구
 소 로봇개발실장