

휴대폰 카메라로 촬영한 악보 영상 인식을 위한 의사트리 알고리즘

Decision-Tree Algorithm for Recognition of Music Score Images Obtained by Mobile Phone Camera

박건희*, 오성열*, 손화정**, 유재명*, 김수형*, 이귀상*
전남대학교 공과대학 전산학과*, 삼성전자 디지털 미디어 사업부**

Keon-Hee Park(tenor12@naver.com)*, Sung-Ryul Oh(acecap@hanmail.net)*,
Hwa-Jeong Son(hwajeong.son@samsung.com)**, Jae-Myeong Yoo(wesleyok@hotmail.com)*,
Soo-Hyung Kim(shkim@chonnam.ac.kr)*, Guee-Sang Lee(gslee@chonnam.ac.kr)*

요약

현대 사회에서 빼놓을 수 없는 기기인 휴대폰 카메라를 통하여 획득한 악보를 인식함으로써 누구나 손쉽게 전문적인 악보에 대한 지식이 없어도 악보를 연주할 수 있는 시스템을 제안한다. 본 실험은 휴대폰 카메라를 이용하여 촬영한 악보 영상을 전처리과정을 통하여 분리된 심볼들을 인식한 후 미디를 구성한다. 본 논문에서는 실험을 위하여 휴대폰 카메라로 촬영한 임의의 악보 영상 11종을 사용하였다. 전처리과정을 거친 심볼을 대상으로 제안한 방법을 통하여 인식한 결과 평균 98%의 높은 인식률을 보였다. 본 시스템을 휴대폰에 포팅하여 수행시간을 측정하고, 영상의 입력 후 미디 생성까지 걸리는 시간이 평균 8.63초가 소요됨을 알 수 있었다.

■ 중심어 : | 악보 영상 인식 | 심볼 인식 |

Abstract

Today, mobile phone is a necessity of modern life. For that reason, we suggest a particular system of a mobile phone which take a picture of music score image and automatically play it without any technical knowledges about the music score information. This experiment makes midi, acknowledging separate symbols via preprocessing to music score image taken. This paper utilizes 11 sorts of the score image taken by a mobile phone camera for this experiment. Through this method we suggest, as much as 98% on average takes place, which is very high recognizing ratio. Also, as we introduce this system in a mobile phone by porting, it takes 8.63 seconds on average to create midi following input of images.

■ keyword : | Music Score Recognition | Symbol Recognition |

I. 서론

최근 컴퓨터 기술이 발달하면서 악보 영상을 자동으

로 인식하여 연주해주는 악보 인식(Music Score Recognition) 시스템에 관한 연구가 활발히 이루어지고 있다[1-9].

* 본 연구는 문화관광부 및 한국문화콘텐츠진흥원이 주관하는 전남대학교 문화 콘텐츠 기술 연구소 육성사업의 연구비 지원에 의해 수행되었습니다.

접수번호 : #071228-001

접수일자 : 2007년 12월 28일

심사완료일 : 2008년 05월 26일

교신저자 : 오성열, e-mail : acecap@hanmail.net

현재까지의 연구를 통해 악보인식을 위한 전체적인 시스템 구성은 정형화 됐다. 조명보정과 오선 검출 및 삭제에 해당하는 전처리 단계가 그 첫 번째로서 이는 입력받는 장치에 종속되는 영상의 질을 인식 가능한 형태로 개선하고 악보의 주요 영역인 오선에 대한 인식 및 삭제에 관한 연구다. 이 연구에서는 수직, 수평 방향의 히스토그램을 이용하여 빈도수를 측정하는 방법, Hough 변환을 이용하여 직선을 추출하는 방법, 검은 화소의 연속적 길이의 빈도수를 측정하여 추출하는 방법, Dynamic Planning 매칭 방법을 사용하여 오선의 연결여부를 추출하는 방법 등이 연구되고 있다[2-5]. 하지만 이들 연구의 공통점은 스캐너를 사용하여 인쇄 악보를 입력영상으로 사용한다는 것이다. 이는 질 좋은 영상을 대상으로 하여 높은 인식률을 갖는다는 장점을 갖지만 스캐너와 같은 부가 장비가 반드시 필요하고 악보의 기울어짐에 대해서 고려하지 못했으며 오선이 휘어지게 인식되는 경우 인식률이 현저하게 떨어지는 한계를 지닌다.

그 다음 과정으로서 각 음표 및 기호의 영역을 정하고 그 음표와 기호를 인식하는 인식단계가 있다. 이 과정은 각 음표의 특성과 기호법의 특징을 이용한다. 박철우 등은 음을 결정하는 조성기호가 악보의 시작부분에 위치한다는 특징을 이용하고 비슷한 음표들의 차영상을 이용하여 음표들을 구분하였다[6]. 또한 사용 빈도수가 높은 음표를 기준으로 패턴 매칭을 이용하여 각 음표들을 구분하는 방법을 사용하는 연구도 있다[7]. 이들 연구는 상당히 진척되어 현재 상용화되고 있는 단계이다. 최근에는 손으로 그린 악보에 대해서도 인식하는 연구가 진행 중에 있다. 하지만 그러는 방법이 매우 제한적이어서 앞으로 많은 개선이 요구되고 있다[8].

악보 영상을 획득할 수 있는 장치로 디지털 카메라나 휴대폰과 같은 휴대가 용이한 장치를 이용하여 악보 영상을 획득하고 이를 연주해주는 악보 영상 인식 시스템을 구현하면, 실용적인 면에서 훨씬 유용할 것으로 판단된다. 따라서 본 논문에서는 휴대폰 카메라로 획득한 악보 영상을 대상으로 한다. 또한 이를 직접 휴대폰에 포팅하여 실시간으로 동작하는 성능을 측정하여 제시한다.

악보 인식 시스템은 [그림 1]과 같이 5단계로 구분된다. 본 논문에서는 인식 단계 및 후처리 단계를 다루고자 한다. 전처리 단계는 음표 및 기호를 인식하기 전에 필요한 모든 과정을 말한다. 인식 단계는 전처리 단계에서 검출한 음표나 기호들을 하나씩 인식해 나가는 과정이다. 마지막, 후처리 단계는 마디선 정보 또는 음악적 특성 등을 이용하여 오인식된 음표나 기호들을 수정하는 과정을 말한다.

본 논문에서는 전처리 단계를 수행한 후 남겨진 음표나 기호들을 인식하기 위해 각 심볼을 수직성분 및 머리의 수에 따라 10개의 그룹으로 나누고 각 그룹 내의 심볼을 인식하는 방법으로 의사트리 알고리즘을 제안한다.

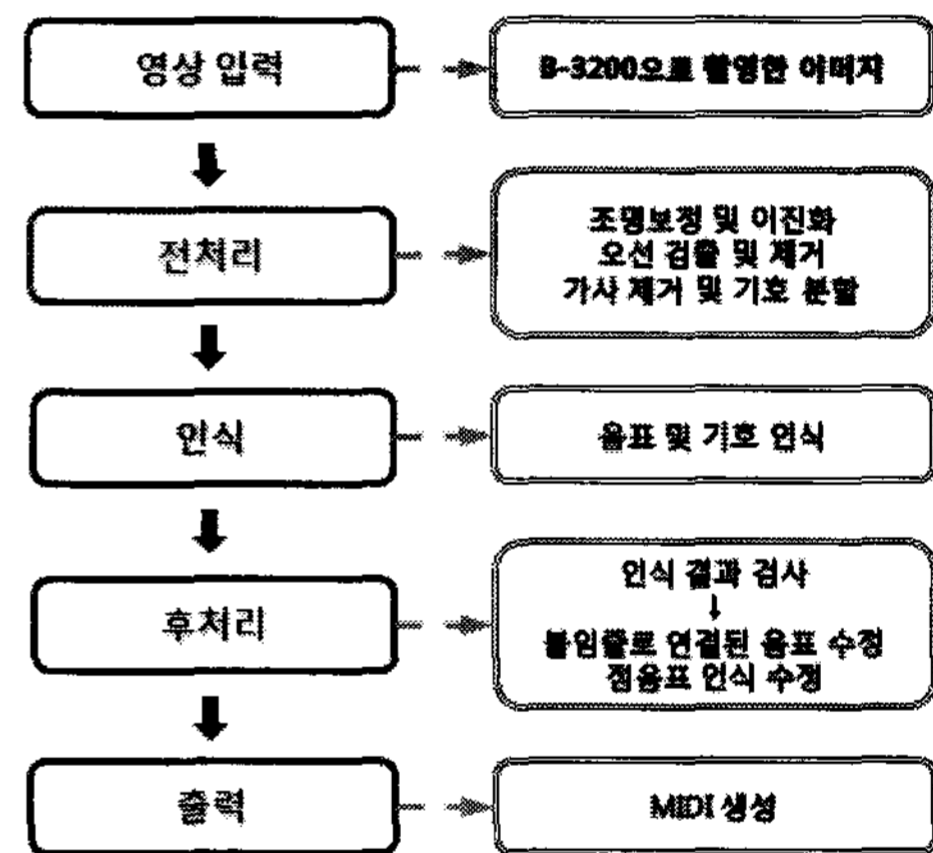


그림 1. 악보 인식 시스템의 구성도

II. 전처리

본 논문에서는 휴대폰 카메라로 촬영한 200 DPI(Dot Per Inch)이하의 저해상도 악보 영상을 대상으로 한다. 휴대폰으로 획득한 영상은 자동 초점이 부정확하여 이미지가 흐릿하여 왜곡된 정보를 포함하고 있다. 또한 임의 각도의 기울어짐이나 휘어짐과 같은 왜곡된 저해상도 영상이다. 이러한 특수한 상황을 고려하여 영상을 이진화하고 기울어짐 정도에 따라 수직 방향으로 분할하고 분할된 영역에 대해서 반복적으로 오선을 검출하고 삭제하는 방법을 사용하였다[9][10].

표 1. 대분류

분류	이름	판별기준	분류기호
1	CLEF	위치	
2	NORMAL_NOTE	$nL=1 \ \& \ nH=1$	
3	BEAM_NOTE	$nL=nH \ \& \ nL \geq 2$	
4	FULL2_NOTE	$nL=1 \ \& \ nH=0$	
5	SHARP_NATURAL	$nL=2 \ \& \ nH=0$	
6	NO_LINE	$nL=0$	
7	MULTI_NOTE	$(nL < nH \ \& \ nL=1)$ or $(nL=2 \ \& \ nH=1)$	
8	MULTI_BEAM_NOTE	$nL < nH \ \& \ nL \geq 2$	
9	MEASURE	위치	
10	CODE	위치	C D E F G A B \flat \sharp m7

III. 특징추출

본 논문에서는 전처리 과정을 통하여 얻어진 심볼을 아래 [표 1]같이 수직성분(nL)과 머리성분(nH) 및 위치 정보에 따라 10개의 그룹으로 대분류한다. 그리고, 각 인식과정의 휴리스틱 값을 갖는 임계치들은 실험 영상을 기준으로 음표의 평균 높이와 머리의 크기를 측정하여 가장 높은 인식률을 갖는 값을 정하였다.

1. 수직성분 검출

전처리 과정을 통하여 분리된 심볼의 수직 방향 히스토그램을 이용하여 수직성분을 검출한다. 심볼의 수직 방향 히스토그램의 최대값을 구한 후 최대값×0.8을 임계값으로 설정한다. 이 임계값이 오선간격×1.5배 이상인 경우에 한하여 수직성분이라고 보고 심볼의 좌측에서 우측으로 탐색해가면서 수직성분의 히스토그램의 값×0.8 이상의 경우에는 동일한 수직선으로 판단하고 넘어간다.

2. 검은 머리 검출 및 판단

검은 머리의 검출은 대분류 2,3,7,8에 해당하는 경우로 앞서 계산한 심볼의 첫 번째 수직선의 위치가 오선

간격 보다 크다면 이 심볼의 머리는 왼쪽에 위치하고 그렇지 않은 경우는 오른쪽에 위치한다고 판단한다. 일반적으로 머리의 두께는 오선 간격과 같다. 본 논문에서는 머리가 2개까지만 있는 경우만을 가정하였다.

심볼의 머리가 왼쪽에 있는 경우를 인식하는 방법은 다음과 같다.

[단계 1] 심볼의 수직성분의 첫 번째 위치 i까지 거리의 1/2의 위치를 머리의 X축 중심이라고 설정한다.

[단계 2] 설정된 X축 중심 값을 이용하여 심볼의 수직방향으로 탐색하며 검은 화소의 수를 계산하고 구해진 검은 화소의 수가 오선간격×1.5배 적은 경우에 머리가 1개 있다고 판단하고 계산된 검은 픽셀의 수×1/2의 위치를 Y축 중심이라고 설정한다.

[단계 3] 만약 단계 2에서 구해진 검은 화소의 수가 오선간격×1.5 보다 크다면 머리가 2개로 보고 구해진 검은 화소의 수×1/4의 위치에 첫 번째 머리가 있는 것으로 판단하고 Y축 중심의 좌표를 계산한다.

[단계 4] 입력받은 중심 좌표(x,y)를 중심으로 사각형 구조요소를 정의하고 임계값 이상의 검은

픽셀이 들어오면 이를 머리로 판단한다.

[단계 5] 처음으로 판단한 머리의 중심 좌표Y로부터 오선간격×0.6 만큼씩 y축 좌표를 더해가면서 단계 4를 반복 수행한다. y축의 좌표가 심볼의 중심위치 보다 작아지면 더 이상 머리가 없는 것으로 판단하고 종료한다.

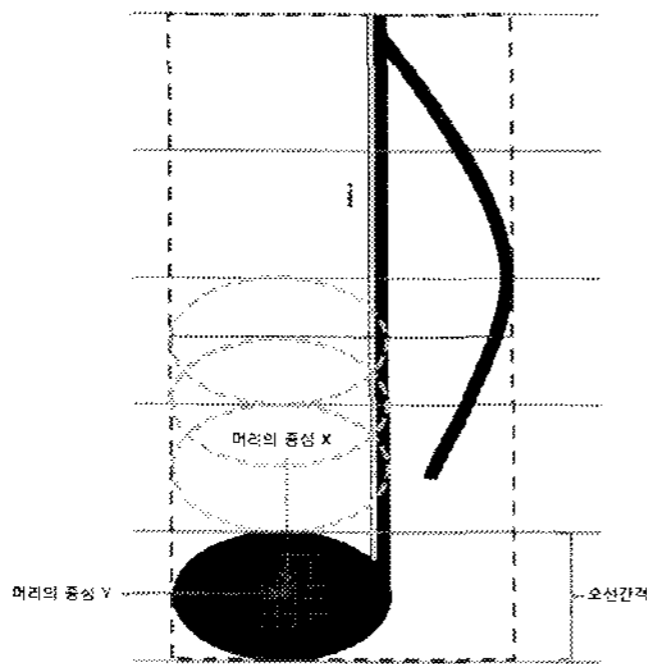


그림 2. 검은 머리가 아래 있는 경우

심볼의 머리가 오른쪽에 있는 경우를 인식하는 방법은 다음과 같다.

[단계 1] 심볼의 첫 번째 수직성분의 위치i로부터 오선간격×0.5 만큼을 머리의 X축 중심이라고 설정한다.

[단계 2] 설정된 X축 중심 값을 이용하여 심볼의 수직 방향으로 심볼의 중심 위치까지 탐색하며 검은 화소의수를 계산하고 머리가 왼쪽에 있는 경우와 같이 Y축 중심의 값을 구한다.

[단계 3] 머리가 왼쪽에 있는 경우의 단계 3과 같다.

[단계 4] 머리가 왼쪽에 있는 경우의 단계 4와 같다.

[단계 5] 처음으로 판단한 머리의 중심 좌표Y로부터 오선간격×0.6 만큼씩 y축 좌표를 감소해가면서 단계 4를 반복 수행한다. y축의 좌표가 심볼의 중심위치 보다 커지면 더 이상 머리가 없는 것으로 판단하고 종료한다.

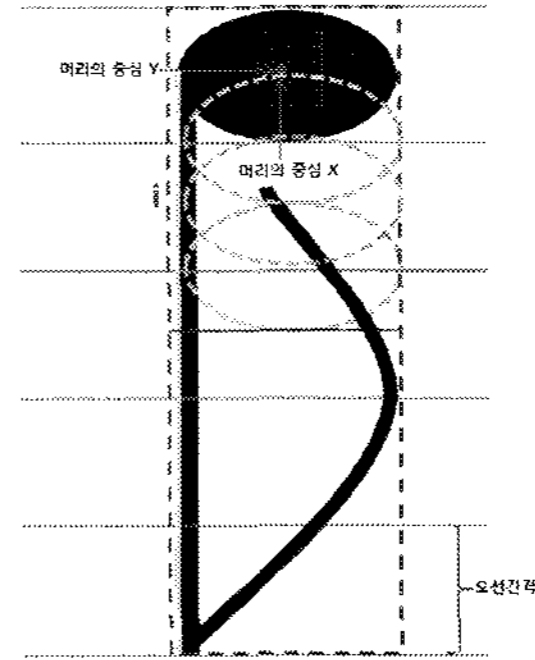


그림 3. 검은 머리가 위쪽에 있는 경우

3. 흰머리 검출 및 판단

흰머리의 검출은 대분류 4로 분류된 경우로 2분음표, 2분음표가 겹음정으로 있는 경우에 해당한다. 머리의 X축 중심은 검은 머리의 X축을 구하는 방법과 같고 Y축의 중심은 오선간격×0.5 만큼이 머리의 Y축 중심이라고 보고 머리의 중심 좌표(x,y)로부터 상·하 및 좌측에 사각형 구조요소를 정의한 후 이 사각형 구조요소에 35% 이상의 검은 픽셀이 있을 경우에 머리가 있다고 판단한다. 겹음정으로 있는 경우에도 검은 머리 검출과 마찬가지로 Y축 좌표를 오선간격×0.6 만큼씩 y축 좌표를 더하거나 빼는 방법으로 다음 머리를 검색한다. 다음 [그림 4]는 흰머리 검출의 예를 보여준다.

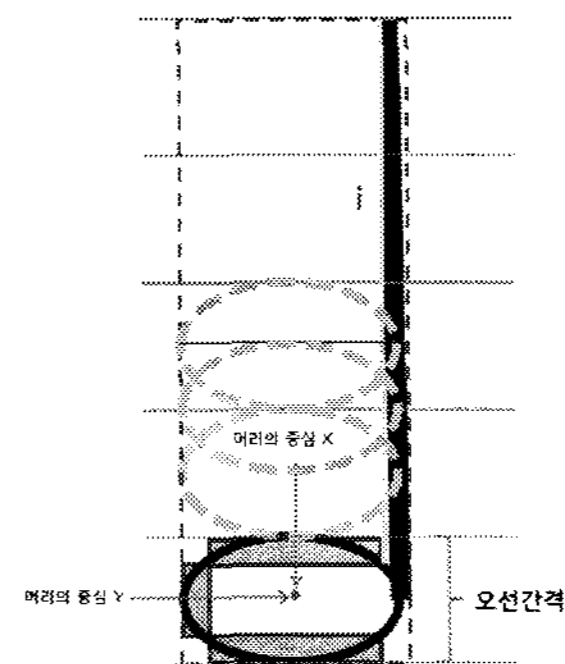


그림 4. 흰머리의 검출

4. 꼬리 성분 인식

심볼의 머리의 위치를 보고 꼬리의 위치를 탐색한다. 머리가 왼쪽에 있는 경우 수직성분의 오른쪽에 있다고

보고 처음 만나는 수직성분 i 의 x 축 위치+오선간격 $\times 0.2$ 의 x 축 위치에서 수직방향으로 오선간격 $\times 2$ 의 위치까지 탐색을 하면서 만나는 검은 화소의 블랍넘버를 보고 꼬리를 판단한다. 수직 방향으로 탐색하며 연속되지 않은 검은 화소를 만나는 경우 처음 만난 꼬리와의 거리차가 4이상인 경우에만 두 번째 꼬리로 인식한다. 다음 [그림 5]는 꼬리의 수를 판단하는 예를 보여준다.

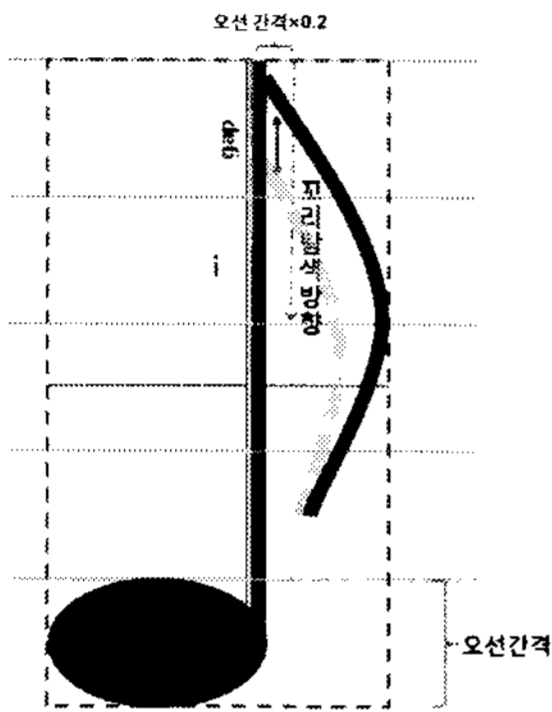


그림 5. 꼬리의 개수 판단

5. 심볼의 오선상 위치 설정

심볼의 오선상 위치는 음표의 경우에 머리의 Y축 중심에 해당하는 위치, 그리고 심볼의 경우는 높이의 중심에 해당하는 위치를 찾는다. 위치를 판단하는 방법은 다음과 같다. 먼저, 심볼의 중심이 오선 단의 위쪽, 오선 단의 사이 또는 오선단의 아래쪽 중 어느 위치에 있는지를 검사한다. 만약 심볼의 중심과 가장 가까이 있는 오선이 첫 번째 오선이고 그 차이가 오선간격 $\times 1/4$ 보다 크면서, 심볼의 위치가 첫 번째 오선의 시작점보다 위쪽에 있으면, 심볼은 오선의 위쪽에 위치한다고 판단한다. 그렇지 않고 심볼의 중심과 가장 가까이 있는 오선이 다섯 번째 오선이면, 그 차이가 오선 간격 $\times 1/4$ 보다 크다면 오선의 아래쪽에 위치해 있다고 판단한다. 위의 두 가지 경우가 아닐 경우에는 오선 내에 위치해 있다고 판단한다. 심볼이 오선의 어느쪽에 위치하는지를 판단한 다음, 중심 위치와 가장 가까이 있는 오선의 위치와 차를 이용하여 다음 [그림 6]에서처럼 어느 위치에 걸쳐 있는지를 숫자로 나타낸다.

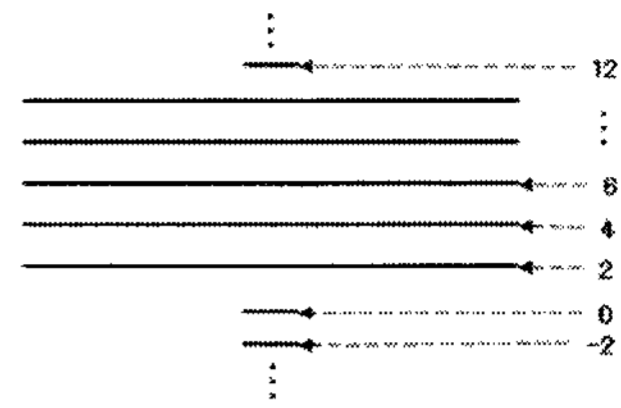


그림 6. 오선상 위치 설정

IV. 대분류

본 논문에서 제안하는 음표인식 알고리즘은 수직성분과 머리성분 및 위치 정보에 따라 10개의 그룹으로 대분류하며 구체적인 분류기준은 다음과 같다.

1. 높은음자리표(CLEF)

음자리표는 높은음자리표와 낮은음자리표로 구분된다. 본 논문에서는 높은음자리표만을 인식한다. 높은음자리표는 1단 악보의 보표의 맨 좌측에 나타나는 심볼이며 그 폭이 다른 기호에 비해 큰 특징을 갖는다. 심볼의 높이가 오선단 높이+오선간격 $\times 2$ 보다 크고, 심볼의 중심의 위치가 오선단의 중앙에 위치하는 심볼을 높은음자리표로 인식하였다.

2. 기본음표(NORMAL_NOTE) 및 단순 겹음표(MULTI_NOTE)

기본 음표는 4분, 8분, 16분 음표의 세 가지로 소분류한다. 이들 음표는 꼬리의 수를 기준으로 구분하고 인식하였다. 기본음표의 인식은 앞에서 설명한 검은머리 및 꼬리를 인식하는 방법에 설명되어 있다.

3. 대로 연결된 기본 음표(BEAM_NOTE) 및 대로 연결된 겹음표(MULTI_NOTE)

대로 연결된 기본 음표는 다음 [그림 7]과 같이 8분, 16분 음표로 구분된다. 기본 음표와 달리 대로 연결된 음표의 머리를 인식하는 방법은 다음과 같다. 머리가 왼쪽에 있는 경우는 첫 번째 수직성분의 가장 위쪽과

아래쪽 좌표를 찾고 수직성분-오선간격×0.6만큼의 위치에 최하위 검은 픽셀의 y축 좌표를 구한 다음 오선간격×0.6 만큼을 올라온 위치가 머리의 Y축 중심이라고 보고 앞서 설명한 검은 머리가 왼쪽에 있는 경우 4-5단계를 수행하여 머리를 판단하였다. 머리가 왼쪽에 있는 경우는 가장 위쪽의 검은 화소의 좌표로부터 오선간격×0.6 만큼 내려온 위치가 첫 번째 검은머리의 중심이라고 보고 앞서 설명한 검은 머리가 왼쪽에 있는 경우 4-5단계를 수행하여 머리를 판단한다.

4. 플랫, 2분 음표, 그리고 마디선(FULL2_NOTE)

플랫은 심볼의 높이를 기준으로 쉽게 구분이 가능하다. 심볼의 높이가 오선간격×2.5보다 작다면 플랫으로 인식한다. 심볼의 너비가 오선간격×1/2보다 작거나 또는 심볼의 검은 화소의 비율이 0.7이상일 경우에 또한 마디선으로 인식한다. 심볼이 플랫 또는 마디선으로 인식되지 않은 경우에는 2분음표로 판단하고 음표의 머리를 검출한다. 앞서 설명한 흰머리를 검출하는 방법에 설명과 같다.

5. 샵, 제자리표(Sharp_Natural)

샵과 제자리표는 수직성분의 시작 좌표의 차이가 오선간격×1/3보다 작거나 같으면 샵으로 인식하고 그렇지 않을 경우 제자리표로 인식한다.

6. 직선 성분이 없는 기호(NO_LINE)

머리와 수직선이 없는 경우의 심볼로 다양한 종류의 음표가 이에 속한다. 다음 그림 은 직선 성분이 없는 심볼의 예를 보여준다. 이러한 심볼을 인식하는 방법으로 점음표의 경우 오선 간격×4/5 이하인 경우에 점음표로 인식하고 후처리 과정을 통하여 앞 심볼과의 관계를 이용하여 잠영인지 점음표인지를 판단한다. 점음표로 인식되지 않은 심볼들 중 기호의 높이가 오선간격×3/4보다 작거나 같으면서 너비가 오선 간격 이상인 경우 온음표와 온십표, 2분 음표중 하나로 판단하고 검은 화소의 비율이 0.7보다 작다면 2분 음표로 인식한다. 그렇지 않은 경우에 심볼의 가장 위쪽과 가장 아래쪽 검은

화소의 개수를 계산하고 위쪽의 검은 화소의 수가 크면 2분 음표, 작으면 온 음표로 구분한다. 나머지 분류되지 않은 경우에는 4분 음표, 8분 음표, 16분 음표중에 하나로 인식하게 되는데, 심볼의 높이의 비에 따라 각각의 음표를 인식하였다.

7. 박자(MEASURE)

박자는 일반적으로 높은음자리표 또는 플랫이나 샵과 같은 조표 다음에 위치한다. 따라서 모든 심볼을 인식한 이후 후처리 과정에서 박자를 인식한다. 인식 대상이 되는 심볼의 앞 심볼이 높은 음자리표인지 아니면 조표인지를 확인하고, 동시에 크기 정보 및 위치 정보를 비교하여 현재 심볼이 박자인지 아닌지를 판단한다. 박자로 인식이 된 경우 심볼을 상위영역과 하위 영역으로 분할하고 숫자를 인식한다. 박자의 상위영역에는 일반적으로 2, 3, 4, 6의 숫자가 나올 수 있고, 하위 영역은 2, 4, 8의 숫자가 나올 수 있다. 분할된 영역을 3×4의 메쉬 특징을 뽑고 미리 정의된 숫자의 메쉬 특징과 최소 거리 분류를 통하여 인식한다.

8. 코드(CODE)

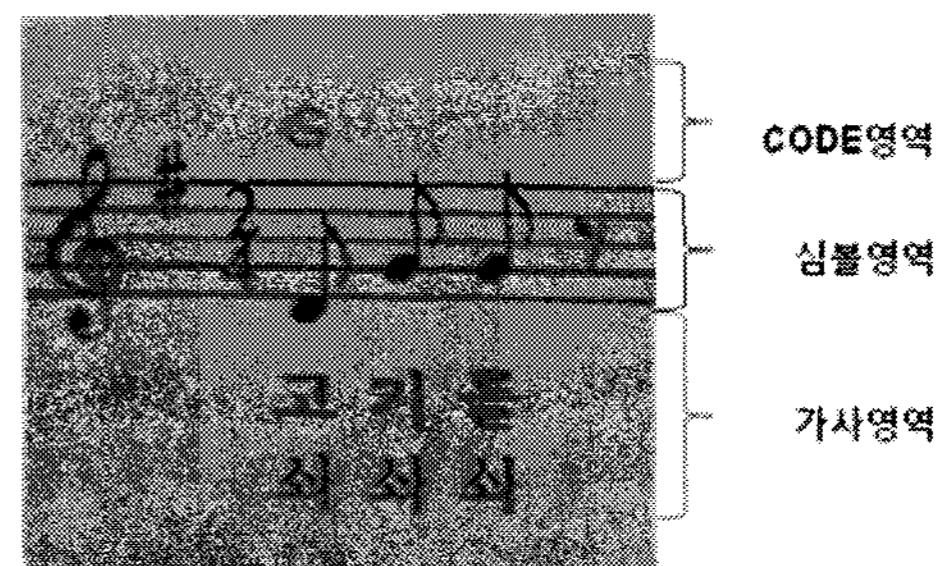


그림 7. 코드 영역 설정

코드는 위 [그림 7]과 같이 악보상에 코드영역, 오선영역, 가사영역을 3부분으로 분할한 다음에 코드영역에 나오는 심볼을 코드라고 보고 인식 대상으로 선정, 심볼의 넓이에 따라 각각 메인 코드와 서브 코드 그리고 인식 대상에서 제외하는 심볼로 분류하였다. 메인코드의 경우 심볼의 넓이가 오선간격×0.9보다 크거나 같은 경우에 해당하며 4×5차원의 메쉬 특징을 이용하고, 서

브코드의 경우는 오선간격×0.9보다 크고 오선간격×오선간격×2 보다 작은 경우이며 3*4의 메쉬 특징을 이용한다. 기타 경우는 인식 대상에서 제외한다. 미리 정의해둔 코드의 메쉬 특징 값들과 비교 최소 거리 분류를 통하여 코드를 인식하였다. 본 논문에서는 알파벳 7개, 샵, 플랫, 소문자 m, 그리고 숫자 7만을 인식하였다.

V. 후처리

후처리는 모든 인식 과정을 끝난 후 오인식된 심볼을 정인식으로 수정하는 단계를 말한다. 본 논문에서는 후처리 단계에서 점음표 인식, 붙임줄로 연결된 음표 연결, 오인식된 샵, 플랫을 수정 하게 된다. 여기서는 미디 파일을 만들기 전에 하는 후처리는 제외한다.

1. 점음표 후처리

점으로 인식된 심볼의 앞심볼의 대분류가 2,3,7,8 인 경우와 2분 음표가 앞 심볼인 경우에 음표의 머리와 점음표의 위치를 비교하여 높이차가 ±2인 경우만을 점이라고 판단하고 그 외의 점으로 판단된 경우는 잡영이라고 판단하고 삭제한다.

2. 붙임줄로 연결된 음표

붙임줄로 연결된 음표는 2가지 종류가 있다. 음의 높이가 같은 경우와 다른 경우로 나눌 수 있다. 음의 높이가 다른 경우는 연주시에 다른 음을 연주하므로 제외한다. 붙임줄의 시작과 끝을 찾고 시작 위치의 음표와 끝나는 위치의 음표의 높이가 같은 경우 끝에 걸려 있는 음표의 음의 길이를 삭제하고 시작 위치에 나오는 음표의 음의 길이를 2배로 바꿔준다. 다음 [그림 8]은 연결한 음표를 수정한 결과이다.

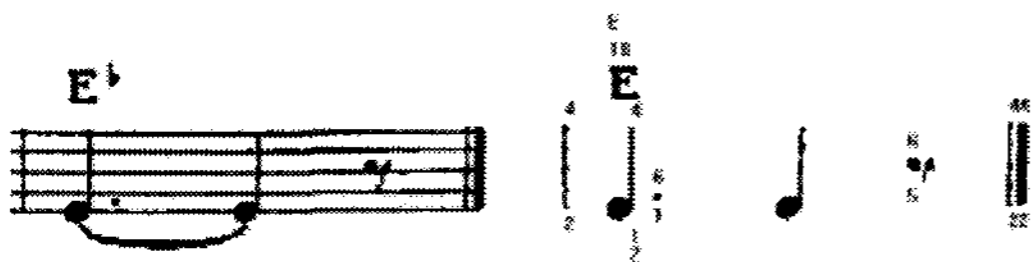


그림 8. 붙임줄로 연결된 음표의 후처리

VI. 인식 결과 출력

지금까지 서술한 방법에 의해 심볼을 인식하고 음표의 경우에는 음의 높이까지를 판단하고 후처리과정을 끝낸 후 화면에 출력되는 인식의 결과를 표기한다. 겹음표는 2개 이상의 음이 겹쳐져 있는 경우를 말하는 것으로, 인식되는 음의 높이는 2개 이상이 된다. 다음 [그림 9]는 인식 결과를 화면에 출력한 예이다.

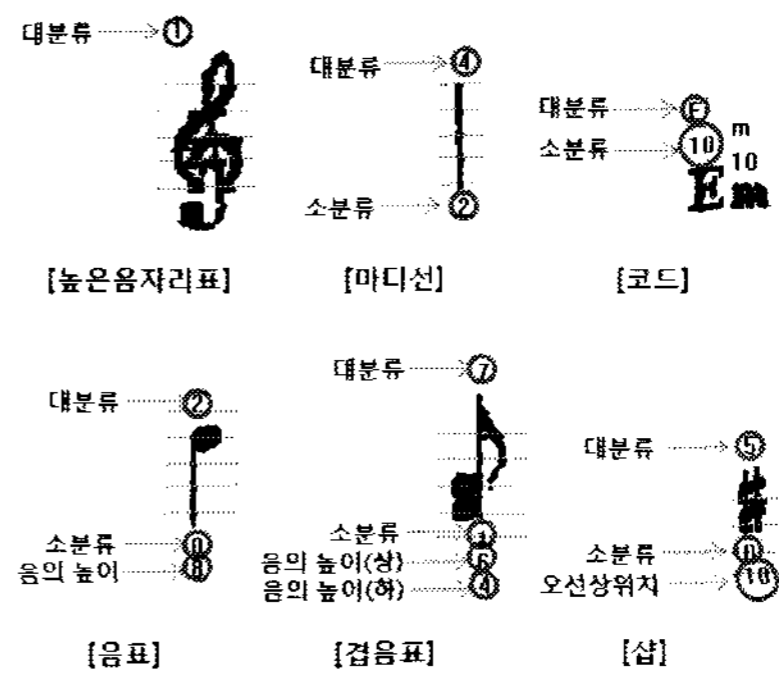
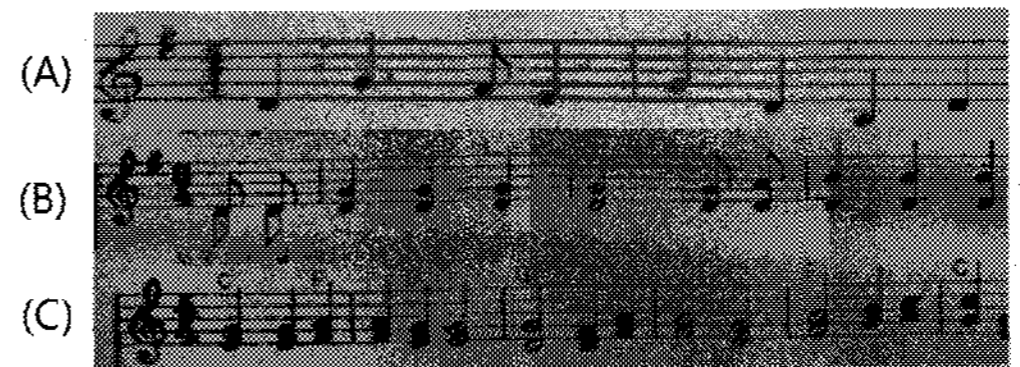


그림 9. 인식 결과의 표기 방법

VII. 미디 출력

MIDI file은 동일 또는 다른 컴퓨터상의 서로 다른 음악 프로그램에서 시간정보를 포함하는 Midi Data를 상호간에 서로 이용할 수 있도록 하는 것으로서 본 논문에서는 Standard Midi format을 사용하여 다음 세가지 악보의 형태와 같이 Midi의 생성형태를 정한다.

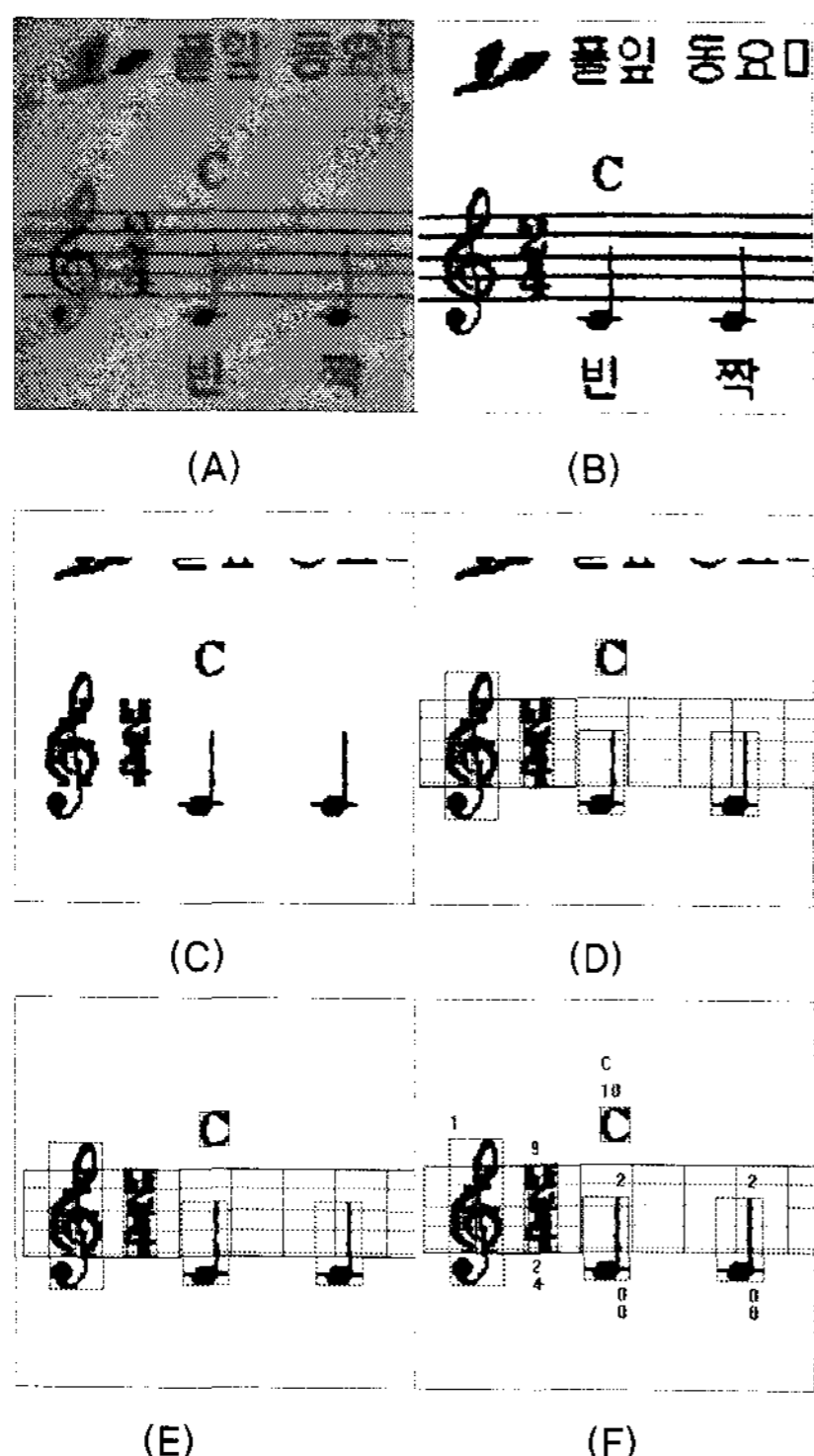


(A)멜로디가 단음적인 경우 (B)멜로디가 겹음적인 경우 (C)멜로디와 코드가 있는 경우

그림 10. 미디 연주 방법

VIII. 실험 및 결과

본 논문은 삼성 휴대폰 B-3200의 자동 초점조절을 지원하는 카메라를 통하여 영상을 입력 받고 심볼의 수직성분 및 머리 성분, 위치로 대분류 한 뒤 분류된 심볼을 의사 결정 기반으로 인식 하여 연주할 수 있는 시스템의 구현하였다. 제안 방법의 인식 성능을 평가하기 위해, 임의의 악보 영상 11종을 촬영하여 실험하였다. 실험 데이터로 쓰인 악보영상은 인식의 난이도에 따라 다음 세가지로 구분할 수 있다. 첫 번째는 가장 기본적인 단계로서 동요수준의 멜로디만 표시된 1단 악보인 애국가와 개구리가 있다, 다음 단계는 단음정 멜로디 외에 기타코드를 인식하여 연주할 수 있게 한 도레미파, 작은별, 고기잡이, 고드름, 개구리 유치원, 가을밤, 꽃밭에서이며, 마지막 단계로는 겹음정 멜로디와 기타 코드로 구성된 악보인 구슬비, 작별이 있다. 다음 [그림 11]은 인식 단계별 결과 영상을 보여준다.



(A) 입력 영상 (B) 이진화 후 결과 영상 (C) 오션 검출 및 삭제 후 영상 (D) 심볼 분리 후 영상 (E) 가사 및 촬영 제거 후 영상 (F) 심볼 인식 후 영상

그림 11. 인식 단계별 결과 영상

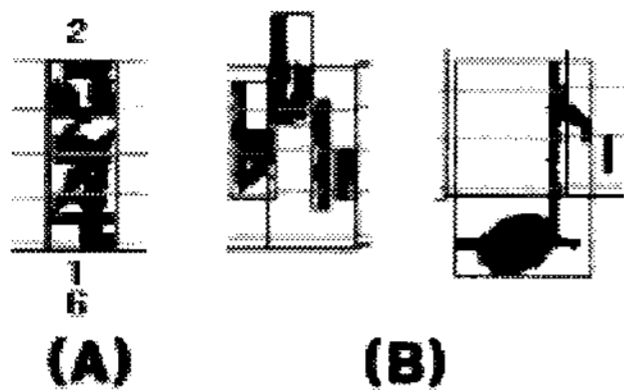
인식의 대상은 악보에 나타나는 여러 가지 보표 중에서 중요한 요소인 음자리표, 박자표, 마디선, 코드, 음표, 쉼표, 조성표이다. 인식률은 오인식 심볼 수를 전체 인식 대상 심볼 수로 나눠서 구했으며, 전처리를 거쳐 오션을 삭제한 후 심볼을 대상으로 측정된 결과 평균 98.2%의 인식률을 얻을 수 있었다. 또한, 본 시스템을 삼성 휴대폰 M-4500에 포팅하여 시스템의 수행 속도를 측정된 결과 상용프로그램으로서 가치를 지닐 수 있는 시간인 평균 8.63초의 수행시간을 보여주었다. 본 실험은 휴대폰 카메라의 특성상 오션의 기울어짐, 오션의 휘어짐, 비네팅 현상이 있는 130만화소 저해상도 영상에서 실험하였다, 또한 기존 PC환경이 아닌, 연산능력이 현저히 떨어지는 휴대폰을 사용하였기 때문에 기존 연구와 성능비교가 불가능하다. 따라서 독립적인 실험 결과를 제시하기로 한다. 다음 [표 2]는 각각의 영상에 대한 인식의 결과를 보여준다.

표 2. 인식 결과

제목	심볼 수	오인식 심볼 수	인식률 (%)
애국가	82	0	100
개구리	103	0	100
도레미파	75	1	98.7
작은별	92	2	97.8
고기잡이	88	1	98.9
고드름	69	2	97.1
개구리유치원	75	3	96
가을밤	107	3	97.2
꽃밭에서	170	5	97.1
구슬비	80	0	100
작별	111	3	97.3
평균인식률			98.2

표에서 볼 수 있듯이 평균 98% 이상의 인식률을 나타내고 있다. 오인식을 일으키는 요인은 연결된 문자가 음표의 대로 인식돼서 음표로 인식하는 경우와 휴대폰 카메라의 특성상 이미지 자체의 손실이 커서 심볼이 분리되는 경우가 있다. 향후 연구에서 문자의 패턴을 정확히 분리하는 부분과 이미지의 복원 알고리즘을 추가하여 개선시킬 수 있을 것이다. 다음 [그림 12]는 오인

식의 예를 보여준다.



(A) 문자를 음표로 오인식 (B) 연결된 요소를 분리함
 그림 12. 인식 오류의 예

IX. 결론 및 고찰

본 논문에서는 휴대폰 카메라로 촬영한 악보 영상을 자동으로 인식하는 방법을 제안하였다. 악보 영상을 자동으로 인식하기 위하여 오선을 검출하고 기울어짐 정도에 따라 서로 다른 개수로 보표를 분할, 분할된 각 부분에 대해서 오선을 검출하였다. 보표를 분할하는 것은 기울어짐이나 휘어짐이 거의 없도록 하는 것으로 별도의 기울어짐 보정 과정이 필요하지 않다. 본 논문에서 제시한 오선 검출 및 삭제에 대한 알고리즘은 [9]의 논문을 인용하였다.

인식의 과정으로 오선을 삭제한 후 분할된 심볼들을 심볼이 가지는 구조적 요소인 수직성분과 머리의 수 또는 위치정보를 이용하여 인식함으로써 매우 빠르고 정확한 인식 결과를 얻을 수 있음을 확인할 수 있었다.

본 시스템을 휴대폰과 같이 휴대가 편리한 이동식 기기에 적용하여 사용한다면 악보와 악기에 대한 음악적인 지식이 없는 일반인들도 쉽게 음악을 연주할 수 있고 감상할 수 있을 것이다. 제안 방법은 현대 사회에서 누구나 가지고 있는 휴대폰 카메라를 이용한 것으로 그 활용 가치가 매우 높다고 할 수 있다.

참고 문헌

[1] D. Bainbridge and T. C. Bell, "The Challenge of Optical Music Recognition," *Computers and the*

Humanities, Vol.35. No.2, pp.95-121, 2001.

[2] I. Fujinaga, "Optical Music Recognition Using Projections," Master's thesis, McGill University, Faculty of Music, Montreal, 1988.

[3] A. Argentini, C. Combi, G. L. Foresti, C. Mirolo, A. Montanari, and A. Peron, "A Multimedia System for Automatic Recognition, Indexing and Retrieval of Heterogeneous Music Documents," *Proceedings of 13th Colloquium on Musical Informatics*, Vol.2, No.5, pp.133-136, 2000.

[4] H. Kato and S. Inokuchi, "A Recognition System for Printed Piano Music Using Musical Knowledge and Constraints," *Structured Document Image Analysis*, H. Baird, H. Bunke and K. Yamamoto, Springer-Verlag, pp.435-455, 1992.

[5] H. Miyao and M. Okamoto, "Stave Extraction for Printed Music Scores Using DP Matching," *Journal of Advanced Computational Intelligence and Intelligent Informatics*, Vol.8, No.2, pp.208-215, 2004.

[6] 박철우, 황인성, 정동석, "인쇄 악보의 자동 인식 및 연주를 위한 연구", 대한전자공학회 추계종합 학술대회논문집, 제16권, 제2호, pp.653-656, 1993.

[7] 황영길, 김성천, "인쇄 악보의 인식과 병렬 알고리즘에 관한 연구", 한국통신학회논문집, 제19권, 제5호, pp.959-970, 1994.

[8] H. Miyao, Minoru, and Maruyama, "An online handwritten music symbol recognition system," *International Journal on Document Analysis and Recognition*, Vol.9, No.1, pp.49-58, 2007.

[9] 손화정, 김수형, 오성열, "카메라 기반 악보 영상 인식을 위한 오선 검출 및 삭제 알고리즘", 한국콘텐츠학회논문지, 제7권, 제11호, pp.34-42, 2007.

[10] 유재명, T. Nguyen, 이귀상, "인쇄 악보 인식을 위해 적응적인 가중치를 이용한 효과적인 이진화 방법 연구", 정보통신분야 합동학술대회 논문집,

