

# 농구 전광판 동영상에서의 의사결정 트리 기반 숫자 인식

은현준\*, 이지원\*, 문성원\*,  
 이정수\*, 남도원\*, 정찬호\*,  
 김창익\*\*

## Decision Tree-Based Digit Recognition in Basketball Scoreboard Videos

Hyunjun Eun\*, Jiwon Lee\*, Sungwon Moon\*,  
 Jungsoo Lee\*, Do-won Nam\*, Chanho Jung\*,  
 Changick Kim\*\*

### 요약

본 논문에서는 농구 전광판 동영상에서의 의사결정 트리를 이용한 숫자 인식 방법을 제안한다. 전광판 내용은 많은 부분이 숫자로 구성되며, 스포츠 경기 분석에 큰 도움을 준다. 제안하는 방법은 농구 전광판에서 팀 점수, 경기 시간 및 쿼터, 공격 제한 시간 인식을 수행한다. 제안하는 방법의 성능 평가를 위해 전광판의 다양한 영역에서 11,269장의 영상을 수집 하였으며, 실험 결과 제안하는 방법이 높은 성능을 가짐을 보였다. 본 연구 결과는 영상기반 숫자 인식 연구, 스포츠 전광판 인식에서 실질적인 도움이 될 것으로 판단된다.

**Key Words** : Basketball game videos, Digit recognition, Decision tree, Sports analysis.

### ABSTRACT

In this paper, we propose a digit recognition

method using a decision tree in basketball scoreboard videos. Many of contents of the scoreboard include digits which are important to analyze on sport videos. The proposed method recognizes team scores, a quarter, a play clock, and a shot clock. For quantitative evaluation, we collected 11,269 digit images. The proposed method achieved high performance on the dataset. We believe that this study guides for digit recognition and scoreboard recognition.

### I. 서론

최근 지능형 스포츠 동영상 분석 시스템을 통한 스포츠 분석 및 이해에 관한 연구가 활발하게 진행되고 있다<sup>[1-6]</sup>. 특히 많은 사람들에게 인기가 높은 축구와 농구 경기를 지능적으로 분석하는 시스템은 팀 및 선수 퍼포먼스 분석, 전략 분석, 심판 판정 보조 등의 다양한 역할을 할 수 있다. 또한 농구 경기의 경우 전광판 내용이 많은 정보를 포함하고 있으며 이를 인식하는 것 역시 경기 분석에 큰 도움을 줄 수 있다. 예를 들어 경기 시간, 경기 쿼터, 공격 제한 시간, 팀 명, 파울 및 점수, 선수 등번호, 이름, 파울 및 점수가 전광판에 나타난다 (그림 1).

본 논문에서는 농구 경기 분석을 위해 전광판 동영상에서 숫자 인식을 위하여 다양한 방법<sup>[7]</sup> 중 의사결정 트리<sup>[8]</sup>를 이용한 방법을 제안한다. 또한 제안하는 방법을 통해 주 전광판에서의 팀 점수, 경기 시간 및



그림 1. 주 전광판 (270×120)과 보조 전광판 (80×60)의 예.  
 Fig. 1. Example of main scoreboard and sub scoreboard.

\* 본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 2016년도 문화기술 연구개발 지원 사업으로 수행 되었습니다 (R2016030044, 스포츠 영상 콘텐츠의 내용 이해 기반 분석/요약/검색 기술 개발).

• First Author : School of Electrical Engineering, Korea Advanced Institute of Science and Technology, hj.eun@kaist.ac.kr, 학생회원

◦ Corresponding Author : Department of Electrical Engineering, Hanbat National University, peterjung@hanbat.ac.kr, 정회원

\* SW·Content Research Laboratory, Electronics and Telecommunications Research Institute (ETRI), {ez1005, moonstarry, jslee2365, dwnam}@etri.re.kr, 정회원

\*\* School of Electrical Engineering, Korea Advanced Institute of Science and Technology, changick@kaist.ac.kr

논문번호 : KICS2017-09-273, Received September 28, 2014; Revised October 16, 2017; Accepted October 16, 2017

쿼터, 보조 전광판에서의 공격 제한 시간 (그림 1. 파란글씨) 인식을 수행 한다. 먼저 각 영역을 전광판으로부터 분할한 후 모든 영역들에서 숫자를 한자리 씩 분할한다. 다음으로 의사결정 트리를 이용하여 숫자 인식을 수행한다. 제안하는 방법의 정량적인 성능 평가를 위해 전광판에서 11,269장의 숫자 영상을 취득해 데이터 셋으로 구성하였으며, 전체 평균 0.9813의 높은 인식 정확도를 나타냈다.

## II. 제안하는 방법

입력은 농구 전광판 동영상 프레임  $B_k$ 의 집합  $V = \{(B_1^M, B_1^S), \dots, (B_K^M, B_K^S)\}$ 으로 정의한다. 여기서  $B_k^M$ 과  $B_k^S$ 는  $k$ 번째 프레임의 주 전광판과 보조 전광판을 의미하며,  $K$ 는 전체 프레임 수를 나타낸다.

### 2.1 인식 영역 분할

먼저 현재 프레임의  $B_k^M$ 과  $B_k^S$ 로부터 경기 시간 영역  $I_k^{RT}$ , 경기 쿼터 영역  $I_k^{GQ}$ , 공격 제한 시간 영역  $I_k^{LT}$ , 어웨이 팀 점수 영역  $I_k^{AS}$ , 홈 팀 점수 영역  $I_k^{HS}$ 을 분할한다. 각 영역은 항상 전광판의 동일한 위치에 존재하기 때문에 고정된 좌표를 이용하여 대략적인 분할이 가능하다. 그림 2에 나타나듯이  $I_k^{GQ}$ ,  $I_k^{LT}$ ,  $I_k^{AS}$ ,  $I_k^{HS}$ 의 경우 해당 영역 숫자의 RGB값에 따라 이진화 후 숫자 크기에 맞게 잘라낸다.



그림 2. 인식 영역 분할.  
Fig. 2. Segmentation of recognition regions.

### 2.2 각 영역의 세부 분할

#### 2.2.1 경기 시간 영역 $I_k^{RT}$

경기 시간의 경우 그림 3과 같이 시간에 따라 4가지로 분류 할 수 있으며, 최대로 나타날 수 있는 숫자는 4개이다. 분과 초 또는 초와 프레임을 구분 짓는 기호를 기준으로 좌측과 우측으로 고정된 크기  $W$ 에



그림 3.  $I_k^{RT}$ 에 대한 세부 분할.  
Fig. 3. Sub-segmentation on  $I_k^{RT}$ .

따라 영역을 분할함으로써 한자리씩 분할이 가능하다. 구분 기호의 경우 커넥티드 컴포넌트를 이용하여 크기가 가장 작은 영역으로 위치 (그림 3. 빨간선)를 정의한다. 본 논문에서는  $70 \times 25$  크기를  $210 \times 75$ 로 리사이징 한  $I_k^{RT}$  대해  $W$ 를 18로 정의 하였다.

#### 2.2.2 공격 제한 시간 영역 $I_k^{LT}$

공격 제한 시간의 경우 그림 4와 3가지로 구분한다.  $I_k^{LT}$ 의 너비에 따라 2자리와 1자리로 분류 할 수 있다. 2자리인 경우 반으로 나누어 각 자릿수를 분할 할 수 있다.



그림 4.  $I_k^{LT}$ 에 대한 세부 분할.  
Fig. 4. Sub-segmentation on  $I_k^{RT}$ .

#### 2.2.3 팀 점수 영역 $I_k^{AS}, I_k^{HS}$

$I_k^{AS}$ 와  $I_k^{HS}$ 의 경우 세 자리까지 나타 날 수 있으며, 한 자리씩 분할이 필요하다. 분할 위치  $x_s$ 는 아래 식으로 정의 된다.

$$x_s = \arg \min_x \left\{ \left| \frac{1}{2} W - x \right| \times \sum_{y=1}^H I_k^T(x, y) \right\} \quad (1)$$

$H$ 는 팀 점수 영역의 높이를 나타낸다.  $W$ 는 두 자리수인 경우 팀 점수 영역의 너비로 정의하며, 세 자리수인 경우 2개의  $x_s$ 를 구해야 함으로 그림 5와 같이 전체 너비의 1/2로 정의 한다. 식 (1)을 이용하여 중심과 가까우며 숫자를 포함하는 픽셀이 가장 적은 위치를 얻을 수 있다.

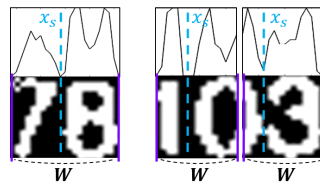


그림 5.  $I_k^{AS}$ 와  $I_k^{HS}$ 에 대한 세부 분할.  
Fig. 5. Sub-segmentation on  $I_k^{AS}$  and  $I_k^{HS}$ .

### 2.3 의사결정 트리 기반 숫자 인식

앞에서 한 자리씩 분할한 숫자를 인식하기 위하여

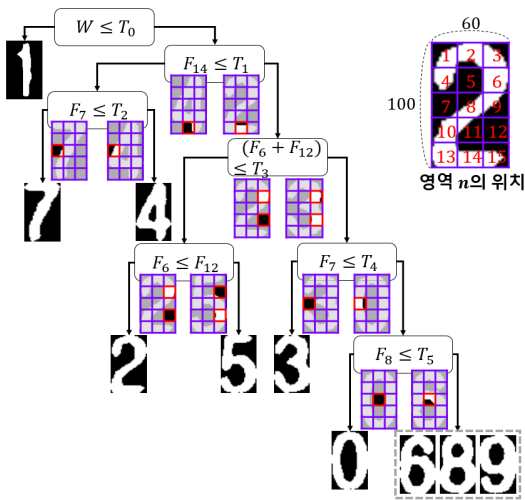


그림 6. 의사결정 트리를 이용한 숫자 인식.  
Fig. 6. Digit recognition using a decision tree.

그림 6과 같이 의사결정 트리 방법을 이용 한다. 1의 경우 이진화된 영상에 기반한 너비  $W$ 가 상대적으로 작으므로  $W \leq T_0$ 인 경우 1로 처음 분류가 가능하다. 다음으로는 영역  $n$ 의 숫자 영역 픽셀 비율  $F_n$ 을 특징으로 분류를 진행한다. 숫자 영역 픽셀 비율은  $F_n = A_n / (h \times w)$ 으로 정의된다.  $h$ 와  $w$ 는 각각 영역  $n$ 의 높이와 너비로 모두 20으로 정의 된다.  $A_n$ 은 이진화 영상에서 영역  $n$ 의 숫자 픽셀의 수이다.  $F_n$ 은 영역  $n$ 에 따라 숫자마다 다르게 나타나기 때문에 숫자를 구분하는 명확한 특징이 될 수 있다. 예를 들어 그림 6에서 나타나듯이 4와 7의 경우 다른 숫자에 비해 7번째 영역의 숫자 픽셀이 적으므로  $F_7$ 을 이용하여 4, 7과 나머지 숫자의 분류가 가능하다.  $T_0, \dots, T_5$ 은 각각 실험적으로 15, 0.5, 0.5, 1.2, 0.5, 0.2로 설정하였다. 파라미터의 경우 일반적으로 해당 영역에 분류하려는 숫자간의 숫자 픽셀 비율 차가 크기 때문에 0.5로 설정할 수 있으며, 그 외의 경우는 실험적으로 최적의 파라미터를 결정하였다. 숫자 6, 8, 9의 경우 그림 7과 같이 좌우측 영역에 대해 인테그랄 프로젝

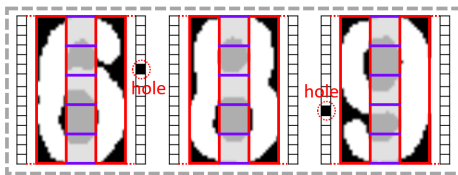


그림 7. 숫자 6, 8, 9의 좌우측 홀 검출.  
Fig. 7. Hole detection for left and right side of six, eight, and nine.

션을 수행하면 위치에 따른 홀의 유무를 알 수 있다. 홀의 유무 및 위치에 따라, 좌측 홀, 홀이 없는 경우, 우측 홀 각각 6, 8, 9로 분류 한다.

### III. 실험 결과

제안하는 방법의 정량적 성능 평가를 위하여 3개의 농구 전광판 동영상에서 영상을 취득 후  $I_k^{RT}, I_k^{LT}, I_k^{AS}, I_k^{HS}$  영역에서 숫자를 분할하여 표 1 과 같이 데이터 셋을 구성하였다. 경기 쿼터의 경우 오랜 시간동안 변화 없이 동일한 숫자를 나타내므로 데이터 셋 취득에서는 제외 하였다. 경기 시간, 공격 제한 시간, 팀 점수 영역에서 각각 6,992, 3,069, 1,208장의 영상을 얻어 전체 10개 클래스 11,269장에 대해 성능

표 1. 데이터 셋을 구성하는 각 영역별 클래스의 수  
Table 1. The number of digits in dataset

| 영역 클래스 | 경기 시간 | 공격 제한 시간 | 팀 점수  | 총 합    |
|--------|-------|----------|-------|--------|
| 0      | 810   | 104      | 119   | 1,033  |
| 1      | 986   | 786      | 321   | 2,093  |
| 2      | 825   | 920      | 80    | 1,825  |
| 3      | 553   | 121      | 56    | 730    |
| 4      | 1,017 | 824      | 108   | 1,949  |
| 5      | 1,103 | 53       | 64    | 1,220  |
| 6      | 431   | 52       | 78    | 561    |
| 7      | 373   | 48       | 118   | 539    |
| 8      | 506   | 75       | 60    | 641    |
| 9      | 388   | 86       | 204   | 678    |
| 총 합    | 6,992 | 3,069    | 1,208 | 11,269 |

표 2. 제안하는 방법의 인식 정확도에 대한 정량적 성능 평가.  
Table 2. Quantitative evaluation on recognition accuracy of the proposed method.

| 영역 클래스 | 경기 시간  | 공격 제한 시간 | 팀 점수   | 평균     |
|--------|--------|----------|--------|--------|
| 0      | 0.9988 | 1.0000   | 1.0000 | 0.9996 |
| 1      | 0.9990 | 0.9975   | 0.9533 | 0.9833 |
| 2      | 0.9915 | 1.0000   | 1.0000 | 0.9972 |
| 3      | 0.9982 | 0.9091   | 1.0000 | 0.9691 |
| 4      | 1.0000 | 1.0000   | 1.0000 | 1.0000 |
| 5      | 1.0000 | 0.9623   | 0.9688 | 0.9770 |
| 6      | 0.8167 | 1.0000   | 1.0000 | 0.9389 |
| 7      | 0.9920 | 1.0000   | 0.9746 | 0.9889 |
| 8      | 0.9704 | 0.9733   | 1.0000 | 0.9812 |
| 9      | 0.9536 | 1.0000   | 0.9804 | 0.9780 |
| 평균     | 0.9720 | 0.9842   | 0.9877 | 0.9813 |

평가를 수행하였다. 표 2에 나타나듯 경기 시간, 공격 제한 시간, 팀 점수의 숫자 인식에서 10개 클래스 평균 0.9720, 0.9842, 0.9877의 정확도를 얻었으며, 각 클래스의 평균 역시 모두 0.93 이상의 높은 인식 정확도를 가진다.

#### IV. 결 론

본 논문에서는 농구 전광판 동영상에서의 의사결정 트리를 이용한 숫자 인식 방법을 제안하였다. 제안하는 방법을 이용하여 경기 시간, 팀 점수, 공격 제한 시간 등을 인식하였으며 10개 숫자 각 클래스 평균 성능이 모두 0.93이상으로 높은 성능을 얻었다. 본 연구 결과는 영상기반 숫자 인식, 스포츠 전광판 인식에 실질적인 도움이 될 것으로 예상된다.

#### References

[1] C. Jung, "Objective evaluation of background subtraction algorithms for soccer video analysis: a experimental comparative study," *J. KICS*, vol. 42, no. 1, pp. 42-45, Jan. 2017.

[2] S. Beysal and P. Duygulu, "Sentioscope: a soccer player tracking system using model field particles," *IEEE Trans. Cir. and Syst. Video Technol.*, vol. 26, no. 7, pp. 1350-1362, Jul. 2016.

[3] C. Jung, "Objective assessment of mathematical morphology operators to improve the accuracy of background subtraction for soccer videos : An experimental comparative study," *J. KICS*, vol. 41, no. 12, pp. 1752-1755, Dec. 2016.

[4] C. Jung, "Impact of image downsampling on the performance of background subtraction in Full-HD soccer videos," *J. KICS*, vol. 42, no. 1, pp. 46-49, Jan. 2017.

[5] J.-H. Anh and J.-H. Choi, "Verification process for stable human detection and tracking," *J. Korea Inst. Inf., Electron., Commun. Technol.*, vol. 4, no. 3, pp. 202-208, Sept. 2011.

[6] M. Merler, B. Huang, L. Xie, G. Hua, and A. Natsev, "Semantic model vectors for complex video event recognition," *IEEE Trans.*

*Multimedia*, vol. 14, no. 1, pp. 88-101, Feb. 2012.

[7] J.-K. Lee, H.-Y. Kim, S.-K. Lee, G.-W. Lee, Y.-R. Park, K.-N. An, C.-S. Bae, and Y.-C. Park, "A study on recognition of both of PCA and LAD using types of vehicle plate," *J. Korea Inst. Inf., Electron., Commun. Technol.*, vol. 6, no. 1, pp. 6-17, Mar. 2013.

[8] B. Pradhan, "A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS," *Computers & Geosciences*, vol. 51, pp. 350-365, Feb. 2013.