

Automatic “Go” Record Generation from a TV Program

Keiji Yanai and Takehisa Hayashiyama

Department of Computer Science, The University of Electro-Communications

E-mail: {yanai, hayash-t}@igo.cs.uec.ac.jp

Abstract

We present a video recognition system of a “Go” TV program. It generates a Go play record automatically from a broadcast of Go played by human professionals. “Go” is the ancient Asian board game played between two player, which is similar to Chess and Shogi. For an MPEG2 video of a TV Go program, the system distinguishes play or commentary board shots from other types of shots such as player’s shots, and detects Go stones placed on the board from board shots. The system removes several types of noise such as a player’s head or hand. In addition, it also detects Go stone from commentary board shots which are often inserted between play board shots, and compensates for the order of Go stones placed on the play board during commentary board shots. In the experimental results for eight TV Go program the system have achieved the 95.7% precision and the 95.7% recall rate.

1. Introduction

Because of recent digitalization of TV broadcasting and wide spread of HDD TV program recorders, the techniques to extract information from TV program for automatic indexing are getting more important. In this paper, we treat with a TV Go program as a case study. We implement a system which generates a Go-play record automatically by recognizing Go stones placed on the board.

Note that “Go” is the ancient Asian game, which is called *baduk* in Korea and *weiqi* in China. Go is one of the most popular board game among China, Korea and Japan. Go is played between two players, Black and White, who alternately place a stone of their own color on an empty intersection on a Go board, with Black playing first. If stones of one color have been completely surrounded by the opponent’s stones, surrounded stones are removed from the board. It is called “capture of stones”. The board size is 19×19 . The goal of the game is to control a larger area than the opponent. Placing of Go stone by a player is called “move”.

Since TV programs of Go play are played between professional Go players in general, their records are sometimes open to the public on books and Web pages. However, such records have just order and places of moves, and they are not associated with movies of Go plays. Therefore, they cannot be used for video indexing tasks such as searching movies for shots when a player make important moves. Moreover, records do not always get open. In this study, we make Go play records associated with movies, which includes information on when moves are made, by recognizing videos of Go plays broadcasted as a TV program in order to enable them be made use of as video indices.

In this paper, we propose a system to generate a Go play record automatically. The system distinguishes shots of a Go board from other shots, and ignores noise such as hands

or heads of players over a Go board. In addition, to improve the precision of generated records, we propose recognizing a commentary board as well as a play board and integrating both recognition results. In general, TV programs have not only shots of a play board where players actually are playing Go but also shots of a commentary board with which commentators explain and analyze the current condition of the game. While commentary shots are shown in the video, some moves might be made on the play board. The order of such moves not shown in the video cannot be recognized by usual video recognition method. So that we integrate results from both boards based on the heuristic that commentary shots can compensate for the order of moves not shown in play board shots since moves on the play board during commentary shots tend to be included in the comments that commentators made and they usually expect or explain them on the commentary board.

In the experiments, we show effectiveness of our proposed method on integration. In fact, the precision of the generated records are improved by introducing the integration method we newly propose in this paper.

As related work, recognition of actual Go and Shogi play have been proposed so far [1][2]. Their systems were designed not for a TV program but for the special setting to capture a Go or Shogi play by a digital camera. So that they did not need to cope with such noise problems that part of player’s body appeared over the board during the play, which make recognition of moves difficult. TV programs are not made for video recognition of moves, so that such situations frequently happen. Therefore, we need to treat with unique problems to TV programs to realize a system which generate Go play records from TV Go programs.

So far, there are a lot of work on video recognition of TV program such as news [3], American football [4], cooking program [5]. However, there has been no work on video recognition of a TV Go program. We believe that this is the first work of Go TV recognition.

2. Structure of a TV Go Program

In this paper, we use a TV Go program, “NHK-cup Go Tournament”, broadcasted by NHK¹ as a target. This program is 100 minutes long for one play. This is a standard-style Go program, so that we can apply the same method to other programs.

Go TV program starts with program title, and introduction of people appearing in the program such as players and commentators. After these opening shots, Go play gets started and is going on, interleaving with commentary shots. There are two types of conditions that Go play is over. One is that one player claims to give up, and the other is that no intersection to place a stone is left. In the latter case, players

¹Japan Broadcasting Corporation



Figure 1. Play board image



Figure 2. Edited play image



Figure 3. Commentary board image



Figure 4. Player's image

arrange stones placed on the board to make counting of both color of stones easier after the play gets over. After the winner is decided, both players discuss and rethink about their play, which are called "Kanso-sen". After that, a TV Go program gets finished. "Kanso-sen" is sometimes omitted from a TV program due to the time limitation. Moreover, play shots are edited as shorter shots to save the time, in case that time of the play is very long.

We give explanation on some types of shots.

1. Play board shots

a. Play board shots (Figure 1)

Mainly, we recognize "moves" from them.

b. Edited play board shots (Figure 2)

Due to the time limitation of TV go program, play board shots are sometimes shortened. In that case, appearance of the shots are slightly different from normal play board shots so that the audience can identify them as edited shots (Figure 2). In our study, we treat with them in the same way as normal play board shots, and we do not distinguish both types of shots.

2. Other shots than play board shots

a. Commentary board shots (Figure 3)

Using a commentary board, commentators expect the next series of moves, explain candidates of next moves and verify actual moves based on the actual play.

b. Player's shots (Figure 4)

The faces of players sometimes appear in the video, while one of the players are thinking about the next move for a little long time. This shots are useless to generate a Go record.

To recognize moves from a TV Go program, there are some problem to be solve. (1) Player's hand or head sometimes occludes play boards. (2) While other shots than play board shots are shown in a program, some moves are made, and the shot at the moment is not available in the video. (3) In the edited shots, moves are made much more quickly than in normal shots. To realize a system to generate a Go record from a TV program, we need to overcome these problems.

3. Algorithm to Generate a Record of Moves

At first, we import the video of a TV program to PC as an MPEG2 file. After that, we carry out the processing as follows for automatic generation of a record of moves.

- (1) Extract one frame image from an MPEG2 file of the Go play video. Classify it into ones of play board shots, ones of commentary board shots, or others. In case of a frame image of play board shots, if a play board has not been detected yet, jump to (2). If detected already, jump to (3). In case of a frame image of commentary board shots, if a commentary board has not been detected yet, jump to (4). If detected already, jump to (5). Otherwise, repeat (1).
- (2) Detect a board from an image of play board shots. If detected, jump to (3). Otherwise, jump back to (1).
- (3) Check the color of pixels around all the intersection, and examine whether a stone has been placed on each intersection. If new stone are detected, jump to (6). Otherwise, jump back to (1).
- (4) Detect a board from an image of commentary board shots. If detected, jump to (6). Otherwise, jump back to (1).
- (5) Check the color of pixels around all the intersection on the commentary board, and examine whether a stone has been placed on each intersection. If new stone are detected, jump to (6). Otherwise, jump back to (1).
- (6) If a new stone, namely a new move, is detected, examine how many frames the stone is detected from. If the stone are detected from more than m frame images, jump to (7). Otherwise, jump to (8).
- (7) Add the detected move to the Go play record.
- (8) Examine whether the play is over or not. If finished, jump to (9). Otherwise, jump back to (1).
- (9) Output the record as a final result.

3.1. Judgment on beginning shot of a Go play

A Go board is not always placed on the exactly same place against the camera, so that we need to detect exact location of a Go board. To judge on beginning of a Go play, we prepare several sample images which are decided as beginning images of a play in advance, and compare each image extracted from video of a TV Go program with sample images. Comparison is carried out as follows:

- (1) Extract a new frame image from the video.
- (2) Compute the difference d between one of sample images and a frame image in terms of RGB color values of all pixels.
- (3) If $d < th$, the frame images are classified as a beginning image, where th is a certain threshold. Otherwise, jump back to (1).

3.2. Classification of a Shot

A TV Go program consists of play board shots, commentary board shots and other kinds of shots. To generate a play record, we need to treat with only play and commentary board shots. To classify a frame image into these three kinds of shots, we prepare sample images of play board and commentary board shots.

As the similar way to detecting the beginning frame, we use difference value between a sample image and a frame image. We use the beginning shots of both play and commentary boards as sample images to classify a frame image. If the difference is small than a certain threshold, we classify it as an image of the shots the difference of which is smaller. Otherwise, we regard it as an image extracted from other kinds of shots.



Figure 5. Go play board just after starting the game

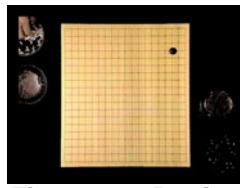


Figure 6. Region detected as Go play board

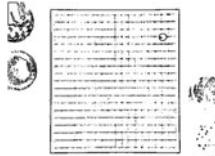


Figure 7. Edge image extracted by Sobel filter



Figure 8. A player's hand is over the board.

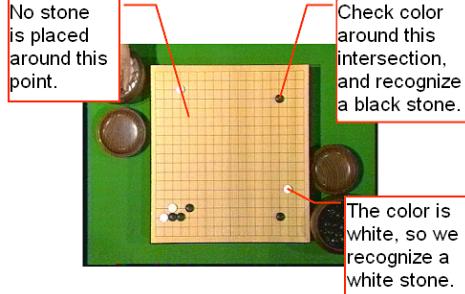


Figure 9. Detecting Go-stones

3.3. Detection of a Play/Commentary Board

We detect a play and commentary board as follows:

- (1) Capture one frame image from the MPEG movie, and carry out "judgment of beginning shots". If it is judged as a frame image of beginning shots, carry out the following processing (Figure 5).
- (2) Extract a region associated with a Go board with the distribution of RGB color value of pixels (Figure 6).
- (3) Detect edges from the image including only the region of a board with Sobel filter (Figure 7).
- (4) Convert the edge image to a binary image.
- (5) Extract vertical and horizontal lines with the Hough transform.
- (6) Estimate 19×19 intersections based on the detected lines, because not all lines of the board are always detected with the above processing due to poor resolution of a frame image captured from a TV broadcasting.

3.4. Detection of Move

We detect a move, namely a newly placed stone on the board, as follows:

- (1) Examine color of pixels around each intersection.
- (2) If pixels of the color of a white or black stone are detected, regard a white or black stone corresponding to the color to be placed on the board (Figure 9).

3.5. Detection of Noise

In case of a TV Go program, a part of player's body such as a hand and a head appears over the Go board. Therefore, we must be able to treat with such situation correctly so that a hand or a head is not detected as a stone by mistake.

To place a stone on the board, a player always use his/her hand, and then the hand always overlaps with the board

(Figure 8). This kinds of images cause recognition error of a move such that a hand is recognized as a white stone or a stone under the hand is recognized to be captured and removed from the board. So that we use not just single frame image, but several frame images to detect one move in order to avoid such mis-recognition. The detail is described as follows:

- (1) Detect new stones on the board. If a new stone is detected, keep on checking if the new stone has been placed for the n consecutive frames.
- (2) After checking the new stone for n frames, if it has been existed for m frames out of n frames at least, regard the new stone as a new move, where $m < n$. In the same way, we detect a captured stone with n consecutive frames.

3.6. Coping with simultaneously detected moves

If moves on the play board are made while other shots than play board shots are shown in the video, multiple moves are detected just after shots are changed to play board shots. And in the edited play board shots moves are sometimes made almost simultaneously, since interval between moves is omitted. In such case, to make a correct record, we have to estimate the order of the moves.

In case that two moves are detected simultaneously, it is not big problem, since we can estimate their order based on the Go rule that both plays make moves alternately. If White player placed a stone just before two moves are detected at the same time, the order of the two missing-order moves can be estimated to be Black first and White next.

In case that more than three moves are detected simultaneously, however, we cannot apply such simple strategy to estimate the order of the missing-order moves. This problem sometimes happen while the commentary board shot are shown in the video. To overcome this hard problem, we propose obtaining information on the order of moves from both play and commentary board shots and integrating them. This is based on the heuristic that commentary shots can compensate for the order of missing-order moves not shown in play board shots since moves on the play board during commentary shots tend to be included in the comments that commentators made and they usually expect or explain them on the commentary board.

We explain the proposed algorithm with a sample case. Two \triangle and one \circlearrowleft are newly detected just after a commentary shot simultaneously shown in Figure 3.6., while we detect the order of moves, $137 \rightarrow 138 \rightarrow 139 \rightarrow 140 \rightarrow 141$, on the commentary board shown in Figure 3.6. Out of these moves, three moves, $137 \rightarrow 138 \rightarrow 141$, were included in the play board just after returning back from the commentary shot, which is shown in Figure 3.6. We compensate for the order of the missing-order moves on the play board with the order of moves appearing on the commentary shot, so that we estimated the correct order of the moves as shown in Figure 3.6.

4. Experimental results

We made experiments with eight TV programs of "49-th NHK-cup Go Tournament" broadcasted on Japan Broadcasting Corporation. These programs are 100 minutes long for one play. In case that the play took more than 100 minutes, the video of the play are edited and shorten so that the program is just 100 minutes long. This is a standard-style

Table 1. Experimental results without integration of detected moves on the commentary boards

kind	number of moves	# of detected moves	# of mis-detected	# detected at same time†	# mis-detected except same‡	recall	precision
Move	2058	2041	113	80	33	93.7%	94.5%
Capture	150	124	11	0	11	75.3%	91.1%

†..the number of more than three simultaneously detected moves out of mis-detected moves

‡..the number of mis-detected moves except simultaneously detected moves

Table 2. Results with integration of moves on both the commentary board and the play board

kind	total moves	detected moves	mis-detected	detected at same(freq.)†	appearing on com. bd. (freq.)‡	detected on com. bd.(freq.)	recall	precision
Move	2058	2041	71	80 (19)	65 (14)	42 (9)	95.7%	95.7%

†..number of more than three simultaneously detected moves out of mis-detected moves

‡..number of mis-detected moves except simultaneously detected moves

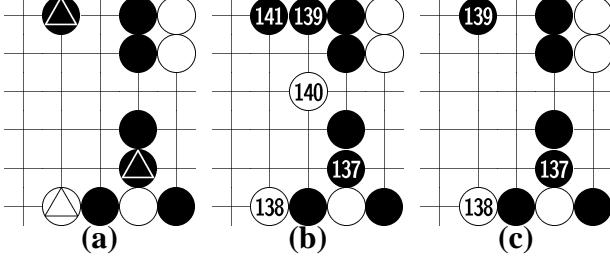


Figure 10. (a)Detected stones on the play board just after the commentary shots
(b)Detected order of moves on the commentary shots (c)the order of moves estimated correctly by the proposed method

Go TV program, so that we can apply the same method to other programs.

4.1. Results without integration of the commentary shots

Table 1 shows the result for eight plays in case of no integration. The recall and the precision are computed by (# correctly detected moves) / (# all actual moves) and (# correctly detected moves) / (# detected moves), respectively.

The recall and the precision for the eight plays in terms of moves were 93.7% and 94.5%. The total number of moves was 2058, which included 113 mis-detected moves and 80 simultaneously detected moves which were missing-order moves. 33 mis-detected moves except simultaneous detection were made mainly due to failure of noise removal such that the player's head and hand were detected as placed stones.

4.2. Results with integration of the commentary shots

We detected moves on the commentary boards and integrate them with the result with no integration described above in order to compensate with the missing order of more than three simultaneously detected moves. Table 2 show the results. In the eight plays, the number of more than three simultaneously detected moves was 80, and 65 moves out of 80 appeared in the commentary board shots. Furthermore, the order of 42 moves out of 65 were reconstructed correctly. As a result, mis-detected moves were reduced by 42 moves, and 71 mis-detected moves remained, which consisted of 15 moves not appearing on the commentary board, 23 moves not detected on the commentary board, and 33 undetected moves resulting from other factors. In the experiments, we reconstructed the order of 7 moves at most, and 4.7 moves on average correctly.

As a final result of integration of both moves from the play and commentary board, we achieved the 95.7% re-

call and the 95.7% precision, which are raised by 2.0% and 1.2%, respectively.

23 moves not detected on the commentary board resulted from frequent overlapping of commentators and the commentary board. 15 more-than-three-simultaneously-detected moves appearing except during the commentary board shots appeared when the interval between moves was very short and just after the shot was very noisy.

5. Conclusion

In this research, we presented the system which generates a play Go record from a TV Go program automatically. To overcome the missing-order move problem, we introduced integrating recognition results of moves on both play and commentary boards based on the heuristic that commentary shots can compensate for the order of moves not shown in play board shots since moves on the play board during commentary shots tend to be included in the comments that commentators made and they usually expect or explain them on the commentary board.

Our experiments showed effectiveness of our proposed method. In fact, the precision and the recall of the generated records were improved by 1.2% and 2.0%, respectively, after integrating the order of moves detected from the commentary board with information on moves from the play boards. Finally we have achieved the 95.7% precision and the 95.7% recall for automatic generation of Go play record associated with the video from the eight TV program in terms of moves.

As future work, we need to improve the recall in terms of capture of stones, and we plan to implement Go video search and abstraction system by making use of generated records as video indices.

References

- [1] D. Kakiuchi, T. Fukuyama, A. Domae, K. Koromo and K. Okazaki: Automatic Recording of Shogi-Record by Image Processing IEICE PRMU report, PRMU2001-149, pp.75-80, 2001.
- [2] T. Shimono, M. Hirano, B. Nen, T. Fukuyama, K. Okazaki and S. Tamura: Automatic Recording of Go-Record by Image Processing 4th SICE Symposium, pp.7-12, 1999.
- [3] S. Satoh, Y. Nakamura, and T. Kanade: Name-It: Naming and Detecting Faces in News Videos, IEEE MultiMedia, Vol. 6, No. 1, January-March, pp.22-35, 1999.
- [4] N. Babaguchi, Y. Kawai, and T. Kitahashi: Event Based Indexing of Broadcasted Sports Video by Intermodal Collaboration, IEEE Trans. Multimedia, Vol.4, No.1, pp.68-75, 2002.
- [5] K. Miura, R. Hamada, I. Ide, S. Sakai, and H. Tanaka: Motion based automatic abstraction of cooking videos, Proc. ACM Multimedia 2002 Workshop on Multimedia Information Retrieval, 2002.