Sports events videos-based semantic content analysis and hidden markov model application

Yongbing Chen¹*, Geng Du²

¹Institute of Physical Education, Wenshan University, Wenshan, (CHINA)
²Sports Training Department, Wuhan Institute of Physical Education, Wuhan, (CHINA)

ABSTRACT

With sports video relative data rapidly development, how to find out required parts of information is focus of current research, the paper proposed sports video semantic analysis method after analyzing hidden Markov model, and proposes Out of play and In play contents analysis method flow, only selects partial segments to test and it gets that the model accuracy has arrived at around 90%, besides it also explains the model can be convenient to lots of information statistics and train as well as other advantages, so the model has universality.

© 2014 Trade Science Inc. - INDIA

KEYWORDS

Hidden markov model;
Semantic content analysis;
Sports events videos;
Mathematical model.

PREFACE

From the seventies, since it had video tape recorder, the way that people watched videos were always playing, fast and suspending as well as other ways; with era development, researches on the item have constantly gone deeper, especially for sports type videos semantic analysis aspect that has become the hotspot of researches in recent years.

Someone has ever researched on the item and put forward his thoughts, such as: Ma proposed to establish a kind of user attention model to analyze sports videos wonderful moment, and the method integrates text characteristics, vision and audition as well as each kind of factors, it tested the segment that might cause audition and vision’s attentions and finally achieved video’s wonderful moment.

On the basis of former research, the paper proposed hidden Markov model, and verifies it by concrete experiments, which proves that the model accuracy is very high, besides it gets the model will play positive roles in the item development in future.

REGARDING HIDDEN MARKOV MODEL ESTABLISHMENT

Markov based on observability cannot be directly applied into practical problems, so the paper introduces hidden Markov, the model is a kind of double random process, it is a series of combination make description by observing that transfers through limited turntable interruption, which can mutual transfer.

In general, the model parameters determination is as following:

Firstly we let symbol set can be expressed by $V = \{v_1, v_2, ..., v_M\}$, and for every state available out-
put number, we can use \( M \) to express; State set in the model, we can use \( S = \{s_1, s_2, \ldots, s_N\} \) to express, the number is \( N \) pieces, for \( t \) moment corresponding state, we can use \( q_t \) to express.

Let factor \( a_{ij} \) to express time \( t \) state \( S_j \), state transition probability distribution matrix is using \( A \) to express, then, \( \pi_{t+1} \) state probability is:

\[
A = \{a_{ij}\}\ \text{where } a_{ij} = P\left[q_{i+1} = S_j, q_i = S_i\right]. \quad 1 \leq i, j \leq N \tag{1}
\]

\( b_j(k) \) represents \( t \) time output symbol \( v_k \) probability, state \( S_j \) observed symbol distribution matrix is using \( B \) to express, and then it has:

\[
B = \{b_j(k)\} \tag{2}
\]

Among them,

\[
b_j(k) = P\left[v_k, \text{int moment } / q_t = S_j\right]. \quad 1 \leq j \leq N, \quad 1 \leq k \leq M \text{, for model probability density, it can make limits, lyric can apply continuous observed density way so at to make consistency estimation on probability density, from which covariance is using } \sum_{jm} \text{ to express, state } j \text{ mixed component the } m \text{ uniformity vector is } \mu_{jm}, \text{ use } \eta \text{ to represent density function, state } j \text{ the } m \text{ mixed coefficient is using } c_{jm} \text{ to express, unknown vector is using } \omega \text{ to express, then finite mixed form is:}
\]

\[
b_j(O) = \sum_{m=1}^{M} c_{jm} \eta(O, \mu_{jm}, \sum_{jm}), \quad 1 \leq j \leq N \tag{3}
\]

When \( t = 1 \), it is using \( \pi \) to express initial state distribution form, that:

\[
\pi = \{\pi_i\}, \pi_i = P\left[q_1 = S_i\right], \quad 1 \leq i \leq N \tag{4}
\]

Adopt Gauss formula \( \eta \) to let \( c_{jm} \) to meet constraint condition as:

\[
\sum_{m=1}^{M} c_{jm} = 1, c_{jm} \geq 0 \quad 1 \leq j \leq N, 1 \leq m \leq M \tag{5}
\]

By above, we can know hidden Markov model involved unknown is:

\[
\lambda = \{A, B, \pi\} \tag{6}
\]

But the model still keeps pace with practice, so it needs to further solve learning, identifying, estimating these three problems so that can apply it into practical problems.

**REGARDING ESTIMATION PROBLEM’S ONE KIND OF FORWARD-BACKWARD ALGORITHMS RESEARCH MODEL**

At first, define that in forward algorithm, forward variable on the condition \( \lambda \), it generated before \( t \) observed symbols partial sequence, \( \{O_1, O_2, \ldots, O_t\} \) and at this time, it is also under \( S_j \) state, then \( \alpha_t(i) \) is:

\[
\alpha_t(i) = P(O_1, O_2, \ldots, O_t; q_t = S_i, \lambda) \tag{7}
\]

Then, following steps are \( P(O / \lambda) \) iterative process:

\( \odot \) Initialize process:

\[
\alpha_1(i) = \pi_i b_i(O_1), \quad 1 \leq i \leq N \tag{8}
\]

\( \odot \) Calculate by iteration:

\[
\alpha_{t+1}(j) = \left[\sum_{i=1}^{N} \alpha_t(i) a_{ij}\right] b_j(O_{t+1}), \quad 1 \leq t \leq T - 1, \quad 1 \leq j \leq N \tag{9}
\]

\( \odot \) It solves the result:

\[
P(O / \lambda) = \sum_{i=1}^{N} \alpha_T(i) \tag{10}
\]

By above forward algorithm, we can deduce backward algorithm, the method reduces computation.

**RESEARCH MODEL BASED ON SPORTS VIDEO SEMANTIC CONTENT**

Regarding earliest applied hidden Markov, it was in the voice identifying, after that, the model was expanded and derived application into statistical timing sequence, it carried out research on sports video semantic content by segmenting and classifying, and regulated top layer is events video that composed of out of play and in play, and it could transfer probability, medium layer was constructed different Markov with different timing relations, and the bottom layer was input layer observed numbers, output was optimal time sequence value, as following Figure 1 show:
On the basis of above provided the model’s frame, we present flow chart about out of play and in play, its figure is as following Figure 2 show:

By above Figure 2, we can get the flow is composed of testing and training two phases, so it needs to establish out of play and in play different topological structures’ hidden Markov model, in the model, \( HMM_O \) and \( HMM_I \) respectively represents out of play and in play hidden Markov model, that:

\[
HMMs = \{ HMM_I, HMM_O \}
\]  

(11)

Among them:

\[
HMM_I = \{ HMM_{I,1}, \ldots, HMM_{I,M} \} \\
HMM_O = \{ HMM_{O,1}, \ldots, HMM_{O,N} \}
\]

Among them, M pieces of different In Play model is expressed by \( HMM_I = \{ HMM_{I,1}, \ldots, HMM_{I,M} \} \), N pieces of different Out Of Play model is expressed by \( HMM_O = \{ HMM_{O,1}, \ldots, HMM_{O,N} \} \), then corresponding state layer is:

\[
S = \{ \text{Loose_view}, \text{Medium_view}, \text{Tight_view} \}
\]

(12)

In the stage, we also adopt forward-backward algorithm and corresponding different hidden Markov probabilities values, and it also needs to take maximum value as node, then:

\[
P(O_t / HMM_I) = \max \{ P(O_t / HMM_{I,1}), \ldots, P(O_t / HMM_{I,M}) \} \\
P(O_t / HMM_O) = \max \{ P(O_t / HMM_{O,1}), \ldots, P(O_t / HMM_{O,N}) \}
\]

Finally by dynamical planning method, it carries out searching so as to achieve optimal interlaced sequence.

Regarding extracted events video features research

At first, due to differences are existing in video’s colors, use Dominant Color Ratio to represent dominant tone ratio values, its abbreviation is \( DCR \), it is the value of whole frame size reciprocal and dominant tone pixel numbers’ product, that:

\[
DCR = \frac{|\text{Pixels of dominant color}|}{|\text{Pixels of the whole frame}|}
\]

(13)

In above formula, frame size is expressed by \( |\text{Pixels of the whole frame}| \) that refers to number of pixels; dominant tone pixel is expressed by \( |\text{Pixels do min ant color}| \).
Except for corresponding colors, we also select movement intense, from which inter-coded macro block number is expressed by $|\text{Inter-coded Macro-blocks}|$, and macro movement direction is expressed by $\sqrt{V_x^2+V_y^2}$, corresponding $y$ and $x$ directions’ movement blocks vectors are expressed by $V_y$ and $V_x$, so the frame movement intense average value is:

$$MI = \frac{1}{\sqrt{V_x^2+V_y^2}} \sum |\text{Inter-coded Macro-blocks}|$$  \hspace{1cm} (14)

By above formula, it can apply MPER into handling with information to estimate movement intense.

### Regarding identification and classification problems research

To look for best state sequence layer, it needs to reveal model’s hidden parts, so it should make classification and identification. By applying above formula different topological structures’ hidden Markov maximum likelihood estimate value, it solves observed vector $O_t$, that:

$$P(O_t | \text{HMM}_t) = \max \left\{ P(O_t | \text{HMM}_{t-1}), \ldots, P(O_t | \text{HMM}_{t-M}) \right\}$$

$$P(O_t | \text{HMM}_t) = \max \left\{ P(O_t | \text{HMM}_{t-1}), \ldots, P(O_t | \text{HMM}_{t-M}) \right\}$$  \hspace{1cm} (15)

Corresponding sequence graph, as following Figure 3 show:

By above formula, it can try to look for best state sequence, in addition, we also need to have the aid of $t$ moment $\psi_t(X)$ array to track $\delta_t(X)$ trajectory changes, from which

$X \in \{\text{In_play, Out_of_play}\}$, so that it can get best scoring state point, so best sequence state process is as following:

I Initialized process

$$\delta_t(\text{In_play}) = P(O_t | \text{HMM}_t)$$

$$\delta_t(\text{Out_of_play}) = P(O_t | \text{HMM}_o)$$

Note: The process doesn’t transfer syntaxes.

II Iterative computational process

Firstly, it should define that $\omega$ represents weight value, it is up to $t$ moment likelihood estimation and syntaxes transferring probability from $y$ to $x$ calculate $\delta_t(X)$ by weighted sum and iteration, then:

$$\psi_t(X) = \arg \max_y \left\{ (1-w)P(O_t | \text{HMM}_x) + w \cdot P_{Y|x} + \delta_{t-1}(Y) \right\}$$

$$= \arg \max_y \left\{ w \cdot P_{Y|x} + \delta_{t-1}(Y) \right\}$$  \hspace{1cm} (16)

Among them:

$$X, \quad Y \in \{\text{In_play, Out_of_play}\},$$

$$2 \leq t \leq T, 0 \leq w \leq 1$$

And:

$$\delta_t(X) = \max_y \left\{ (1-w)P(O_t | \text{HMM}_x) + w \cdot P_{Y|x} + \delta_{t-1}(Y) \right\}$$  \hspace{1cm} (17)

Among them:

$$X, \quad Y \in \{\text{In_Play, Out_of_Play}\}$$

III Best scoring moment is $T$, corresponding best score is $\delta_T(X)$, then state at this time is $q^*_T$, that:

$$q^*_t = \arg \max_{q^*_t} \delta_t(X), \quad X \in \{\text{In_Play, Out_of_Play}\}$$  \hspace{1cm} (18)

IV State sequence recall process:

$$q^*_t = \psi_{t+1}(q^*_t), \quad t = T-1, T-2, \ldots, 1$$  \hspace{1cm} (19)

### Hidden markov model experimental result

By adopting $K$ groups of training values as $O = [O^{(1)}, O^{(2)}; \ldots, O^{(K)}]$, respectively carry out training on Out of Play and In Play, that:

$$a_{ij} = \frac{\sum_{k=1}^{K} \sum_{t=1}^{T-1} \gamma_t^{(k)}(i, j)}{\sum_{k=1}^{K} \sum_{t=1}^{T-1} \gamma_t^{(k)}(i)}$$  \hspace{1cm} (20)
**Figure 4: Tennis and soccer correlation parts**

**TABLE 1: Hidden Markov model-based semantic content analysis experiment data**

<table>
<thead>
<tr>
<th>No.</th>
<th>Events name</th>
<th>Video</th>
<th>Broadcaster</th>
<th>Broadcasting time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tennis1</td>
<td>13 Australian Open men’s singles semifinals segment</td>
<td>9 minutes 36 seconds</td>
<td>CCTV-5</td>
<td>2013-01-23</td>
</tr>
<tr>
<td>Tennis2</td>
<td>13 Australian Open Women’s double finals segment</td>
<td>20 minutes 01 seconds</td>
<td>CCTV-5</td>
<td>2013-01-24</td>
</tr>
<tr>
<td>Tennis3</td>
<td>13 World women’s tennis professional Shanghai segment 1</td>
<td>28 minutes 23 seconds</td>
<td>CCTV-5</td>
<td>2013-09-14</td>
</tr>
<tr>
<td>Tennis4</td>
<td>13 World women’s tennis professional Shanghai segment 2</td>
<td>13 minutes 34 seconds</td>
<td>CCTV-5</td>
<td>2013-09-14</td>
</tr>
<tr>
<td>Tennis5</td>
<td>13 World women’s tennis professional Shanghai segment 3</td>
<td>8 minutes 18 seconds</td>
<td>CCTV-5</td>
<td>2013-09-14</td>
</tr>
<tr>
<td>Soccer1</td>
<td>National women’s united soccer league Tianjin team versus Hebei team segment</td>
<td>18 minutes 43 seconds</td>
<td>CCTV-5</td>
<td>2013-03-21</td>
</tr>
<tr>
<td>Soccer2</td>
<td>First division group A league Tianjin team versus Bayi team segment</td>
<td>15 minutes 43 seconds</td>
<td>CCTV-5</td>
<td>2013-09-21</td>
</tr>
<tr>
<td>Soccer3</td>
<td>Shenzhen Jianlibao team versus Liaoning team segment</td>
<td>7 minutes 59 seconds</td>
<td>CCTV-5</td>
<td>2013-09-14</td>
</tr>
<tr>
<td>Soccer4</td>
<td>FA premier league Manchester United versus Aston villa segment</td>
<td>21 minutes 28 seconds</td>
<td>Hunan Satellite TV</td>
<td>2013-03-15</td>
</tr>
<tr>
<td>Soccer5</td>
<td>FA premier league Fulham versus Manchester United segment</td>
<td>26 minutes 32 seconds</td>
<td>Hunan Satellite TV</td>
<td>2013-03-22</td>
</tr>
</tbody>
</table>

\[
\sum_{i=1}^{K} \sum_{t=1}^{T_i} \gamma_t^{(i)}(j) \quad (21)
\]

\[
\overline{b}_j(k) = \frac{s.t.O_t^{(i)} = v_k}{\sum_{i=1}^{K} \sum_{t=1}^{T_i} \gamma_t^{(i)}(j)} \quad (22)
\]
FULL PAPER

$$\pi = \frac{1}{k} \sum_{k=1}^{K} \gamma_{t}(i)$$  \hspace{1cm} (23)

$$\gamma_{t}(j)$$ and $$\zeta_{t}(i, j)$$ respectively represent in case observed sequence and model are given, $$t$$ moment under $$S_{i}$$ probability and $$S_{j}$$ state $$t + 1$$ moment probability.

Test on above model, its process is firstly solving hidden Markov syntax transfer matrix, and then running maximum likelihood estimation to solve alternating sequence optimal solution, during the process let $$w \in [0, 1, 0.4]$$, we respectively intercept tennis and soccer relative parts wonderful video segments, as following Figure 4 show:

Its experiment process data is as following TABLE 1:

By above Figure 4, we can also research on its accuracy, and list the accuracy concrete values, as following TABLE 2 show:

By above TABLE 2, we can get the model accuracy has nearly arrived at 89%, which proves the model is relative reasonable.

<table>
<thead>
<tr>
<th>Test data</th>
<th>Method in the paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tennis1</td>
<td>0.884</td>
</tr>
<tr>
<td>Tennis2</td>
<td>0.918</td>
</tr>
<tr>
<td>Tennis video</td>
<td></td>
</tr>
<tr>
<td>Tennis3</td>
<td>0.892</td>
</tr>
<tr>
<td>Tennis4</td>
<td>0.868</td>
</tr>
<tr>
<td>Tennis5</td>
<td>0.904</td>
</tr>
<tr>
<td>Average accuracy</td>
<td>0.8928</td>
</tr>
<tr>
<td>Soccer1</td>
<td>0.896</td>
</tr>
<tr>
<td>Soccer2</td>
<td>0.891</td>
</tr>
<tr>
<td>Soccer video</td>
<td></td>
</tr>
<tr>
<td>Soccer3</td>
<td>0.823</td>
</tr>
<tr>
<td>Soccer4</td>
<td>0.861</td>
</tr>
<tr>
<td>Soccer5</td>
<td>0.845</td>
</tr>
<tr>
<td>Average accuracy</td>
<td>0.8836</td>
</tr>
</tbody>
</table>

CONCLUSIONS

The paper targeted sports events videos features, it proposes hidden Markov model, by extracting videos’ colors and movement features to run the algorithm so that it can get optimal sequence, in addition, we analyze accuracy on the model, we can get with the result that the model has higher accuracy, and can promote it to other medias video analysis basis, so it has universalities.

REFERENCES


