MINING ACTOR CORRELATIONS WITH HIERARCHICAL CONCURRENCE PARSING

Kun Yuan, Hongxun Yao, Rongrong Ji, Xiaoshuai Sun
School of Computer Science & Technology, Harbin Institute of Technology, 150001, Harbin, P.R. China
Mail: {kyuan, yhx, rrji, xssun}@vilab.hit.edu.cn Tel: +86+45186416485

ABSTRACT

Mining actor correlations from TV series enables semantic-level video understanding and facilitates users to conduct correlation-based query. In this paper, we introduce a graph-based actor correlations mining framework, which serves as the first attempt for effective actor association presentation and concurrence search. We leverage face detection and tracking to locate actors with 2D-PCA detector as pretreatment. To measure the actor association into a unified graph, we propose a context-based actor correlations hierarchical parsing approach, which considers video structure and hierarchical concurrence to refine actor association in our graph modeling. We not only can accomplish actor correlations mining, but also can acquire higher semantic information according to concurrence change. We present the actor correlation mining results in a graph-based interface to enable efficient users’ navigation and search.

Index Terms—Video Content Analysis, Actor Correlations Analysis, Video Context

1. INTRODUCTION

With the proliferation of digital videos, there are increasing requirements in mining semantics from video content to facilitate users to browse videos. Content-based video search usually involves various kinds of search prototypes, such as querying specific actors [1], story placing [2], and actor actions [3]. In most videos, actor correlations are essential cues to help audiences in understanding video scenario. This issue is especially for films and TV series, in which actor correlations can always include higher-level semantic cues to reveal the story scenarios. We had built a demo system “VisualCor” [10], which presents a graph-based interface to enable efficient users’ concurrence navigation. Furthermore, we can achieve new information with actor correlation changes to facilitate users to browse highlight video shots.

Over the past few decades, many researchers have adopted actor information to analyze video content, such as actor video shots retrieval and cast listing. However, users may not pay close attention to video shots list including one actor rather than actor concurrent relationships in the huge volumes of video data. Inspired by the work from “Person-Guanxi” [4] for person association search, in this paper we present an actor correlations mining method to build relation network based on actor concurrence parsing and visual content analysis techniques. Based on actor correlations mining, we propose a new video retrieval mode: Character Correlations Search, and accomplish further actor correlations analysis, as in Fig. 1.

![Figure 1. Concurrence Graph centered on Chandler in TV series “Friends”. Each line represents concurrence association between two actors, and its length represents how tightly the two actors are related.](image)

![Figure 2. System overview of actor concurrence association method with context information.](image)

Generally speaking, difficulty exists in measuring and unifying actor correlations in different video layers, such as shot-level, scene-level. We fuse “AdaBoost with Cascade”...
face detector [5] and tracker IVT [6] to locate face regions. We adopt context information in sequential shots to mutually understand associations, and present hierarchical concurrence extent schema to calculate concurrent relation extent for ranking video shots. Finally, an actor correlations graph generates for searching and browsing actor relationship in our interface. In addition, we analyze more semantic information by actor correlation changes. The overview of our method is presented in Fig.2.

2. ACTOR INDEXING

We distill actors in videos by firstly carrying out scene and shot detection, then adopt face detection [5] and IVT face tracking [6] to locate actors. We choose 2D-PCA detector to extract facial feature and apply nearest neighbor clustering to index actors with a manual step.

2.1. Locating Actors

Shot boundary detection (SBD) is achieved by a graph partition model [8] [9]. In each shot, we could generate multi-pose face sets by locating actor faces and leveraging face tracking algorithm [6]. After face tracking, a face set may include many face subsets of different poses from the same actor, as in Fig.3 (a).

![Figure 3. (a). Locating face region and multi-pose face sets generation; (b). Actor face sets clustering.](image)

2.2. Clustering Actors

We choose 2D-PCA to extract facial feature with $L^2$ distance measurement. Due to variances pose in face extracting, we discover that features of the same person may distribute discretely in feature space. Given two face sets $F_k$ and $F_l$, pose subsets $p_{F_k}$, $p_{F_l}$, $p_{F_k}$, $p_{F_l}$. The similarity between face set $F_k$ and $F_l$ is calculated using Equation 1 as follows:

$$d(F_k, F_l) = \min_{p_{F_k}, p_{F_l}, p_{F_k}, p_{F_l}}(d(p_{F_k}, p_{F_l}))$$ (1)

We then adopt nearest neighbor and threshold $T$ to cluster faces. If $d(F_k, F_l)$ is the smallest and less than $T$, we recognize that $F_k$ and $F_l$ belong to the same person, and $p_{F_k}$, $p_{F_l}$ are recognized as the same pose, and $F_k$, $F_l$ are integrated into a new face set $S$ as in Fig.2(b). From shot layer to scene layer, we cluster numerous face sets into several actor sets to locate and index actors appearing in videos with a manual step. Indexing unknown actors is the foundation for mining correlations in videos.

3. MINING ACTOR CORRELATIONS

We utilize a context-based actor correlations mining approach, which not only analyze concurrence in single video shots, but also consider context information. Moreover, we leverage hierarchical concurrence to reinforce integral correlations to measure actor correlation and generate actor concurrence graph. According to actor correlation quantization, we analyze actor concurrence changes to acquire highlight video shots list.

3.1. Context-Based Actor Concurrence Graph

After above locating and indexing actors, we achieve $N$ distilled actors. We define context of target shot as its surrounding shot sequence.

3.1.1. Context-based actor concurrence analysis

Two or more actors appear in a shot at the same time, which illustrates that there are close co-occurrence associations among them from the intuitive point of view. However, human-association is very complex in videos, so that analyzing actor appearance in a shot merely is not enough and accurate. A shot and its surrounding shots may represent a plot between two actors in videos. Considering the sitcom story continuity, actor co-occurrence associations are implied in sequential video shots. Therefore, actor correlations for each shot, is analyzed by surrounding shots with a Gaussian weight measurement.

Given a shot sequence $P$ in a scene, and the value of $P^{k}$ denotes whether actor appears in the $k$th shot .

$$\begin{align*}
P^k_i &= 0 & \text{if actor } i \text{ is absent in the } k^{\text{th}} \text{ shot} \\
P^k_i &= 1 & \text{if actor } i \text{ is not}
\end{align*}$$ (2)

If a character $i$ is absent in the $k^{\text{th}}$ shot, which illustrates that there are no co-occurrence associations relevant to $i$ and let $r^k_{ij} = 0$. If not, $P^k_i = 1$. We use a Gaussian weight measurement to calculate relational extent, which bases on video context information and semantic content. As in Fig.4 (left side), to analyze the target shot ($k^*$), we consider actors appearing in k-centered shot sequence within the shot sequences. There is correlation between $i$ and $j$ with Gauss extent, $N(k, \sigma)$, actor $j$ appearing in a shot sequence $P$ given a surrounding span ($k^- \sigma, k^+ \sigma$). We calculate correlations, $r^k_{ij}$, using Equation 4:

$$\begin{align*}
r^k_{ij} &= 0 & P^k_i = 0 \\
r^k_{ij} &= \sum_{i-j} \exp \left( \frac{(n-k)^2}{\sigma^2} \right) \times P^k_i & P^k_i = 1
\end{align*}$$ (3)

The $k^*$ shot is nearer to the $k^*$ one in temporal sequence, which indicates that we should build closer concurrence extent, equal to $r^k_{ij}$. A concurrence matrix representing actor
Figure 4. Proposed actor correlations analysis method Left side: context-based concurrence measurement with Gaussian weight; right side: hierarchical correlations reinforcement in video structure.

Concurrence not only considers context relations but also measures how tightly characters are related.

3.1.2. Hierarchical concurrence analysis

In TV series, main actor correlations exist throughout all the videos as well as video story users concerned. Hence, there are concurrence cues in integral video structure. We propose a hierarchical association refinement method based on video-scene-shot structure in Section 2.

In scene level, we cumulate concurrence matrix \( R \) of each shot in a scene to quantify actor correlations matrix \( SR \), which is defined as follows:

\[
SR(i, j) = \sum_{k=1}^{s_n} R^k(i, j)
\]  

(4)

Where \( SR(i, j) \) is concurrent value between actor \( i \) and \( j \), equal to summation of co-occurrence extent in each shot, and \( s_n \) is the number of shots in scene.

In video level, we not only calculate correlation value, but also measure frequency of correlation appearance in all scenes as a weight to refine results:

\[
VR(i, j) = \left( \sum_{k=1}^{s_n} SR(i, j) \right) \times \frac{S_n}{S}
\]  

(5)

where \( S_n \) is the number of scenes including the correlation between actor \( i \) and actor \( j \), and \( S \) is the number of scenes in videos. In videos, co-occurrence ratio of different scenes indicates extensibility of character correlations. Equation 5 illustrates that more times actor \( i \) and actor \( j \) appearing in various conditions should be of higher extent. Based on the two approaches, we could analyze and calculate correlations among several actors in similar way.

3.2. Ranking Concurrent Shots

By mining actor correlations in Section 3.1, we could index shots including special actor correlations. Our goal is building a graph model for searching users’ more concerned concurrent relations and browsing corresponding shots conveniently. Given a correlation query, we firstly calculate the ranking score \( (\text{RankScore}_{i,j}(k)) \) of correlation in each video shot (the \( k \)th) and then return shot list ordered from high to low using Equation 6 and 7.

\[
\text{RankScore}_{i,j}(k) = R^k(i, i)
\]  

(6)

(2) Character correlations between two persons, \( i \) and \( j \). The results of each top-rank shot lists with the top closer relations are fused together to generate the decision.

\[
\text{RankScore}_{i,j}(k) = R^k(i, j) \quad (i \neq j)
\]  

(7)

3.2. Actor Correlation Changes Analysis

According to above analysis, we could achieve actor concurrence and correlation graph in video database. It is important for users and researchers to mining more new and higher semantic information. Actor correlations analysis acquires concurrence content between two actors. More times two persons appear in the same time, closer they are in shots. In TV series, two actors’ correlation also changes with the changes of story. We analyze the difference of concurrence in different parts of video. If there is a big change between actors’ correlations in two parts of story, we may consider the video shots including correlations as highlight or important video shots. Using Equation 8, we can calculate the change ratio \( H_{lp} \) between two segments in TV series. \( R(i,j)_{A} \) is the correlation measure between actor \( i \) and \( j \) in the part \( A \) in the whole videos.

\[
H_{lp} = \frac{|R(i,j)_{A} - R(i,j)_{B}|}{R_{A}}
\]  

(8)

We could display the analysis results to facilitate users to access and browse videos.

4. EXPERIMENTAL RESULTS

4.1. Experimental Database and Evaluation

In our experiments, we use 20 hours video of “Friends” TV series as our evaluation database. All videos are segmented into single camera shots about 4000 shots. Over 800 face sets are clustered by centered nearest distance into about 50 face sets including main and minor characters by threshold \( T = 0.25 \). Manual labeling is performed on these 60 sets to index actors accurately, based on which we achieve 17 actors in experiment database. We calculate a 17×17 actor concurrency matrix for each shot, then accumulate and refine them to achieve a 17×17 actor’s matrix of all the videos. Based on concurrency matrix, we could achieve values quantified representing actors correlations extent.

4.2. Results and Discussions

We apply Flash technology to present a visual actor correlations graph. A graph-based interface is presented in Fig. 5. In navigation process, users can browse a global actor association graph, such as in Fig. 5(a), and click on an interested actor in which users can understand all actors...
related to their interested one and how tightly actors concurrence are by the linking line length in Fig. 5(b). In search process, users can query a correlation concerned by clicking on the linking line between actors to browse ranking shots list (top 20), such as in Fig. 5(c).

Our experiments are conducted to evaluate the performance of our proposed mining correlations model. The actor concurrence precision in all ranking shots is up to 90% in our video database, and the precision of each two actor’s co-occurrence in ranking top 20 is up to 98%.

Due to variance in face poses, face detection and tracking in a shot may result in face loss and error. However, our approach could distinguish actor concurrence contribution on a target shot from its neighbor shots. This method could identify accurate actor correlation, and the context-based concurrence extent is calculated including video story development information. Secondly, in the hierarchical method, we refine correlation extent in different levels. By adding video structure information in TV series, we consider video integral story structure in concurrence analysis.

5. CONCLUSIONS AND FUTURE WORK

This paper presents a mining actor correlation approach for analyzing actor associations to enable efficient users’ navigation and search. Our work merits in that we propose a new retrieval mode on Character Correlations, and proposes a actor concurrence parsing method in which users can search actor correlations with the return ranking shot list in TV series. Users also can achieve actors’ new information and actor concurrence relations in video story. In our future work, we tend to refine our actor concurrence model by integrating human recognition to achieve better actor clustering results.

Acknowledgments: This work is supported by National Basic Research Program of China (2009CB320906), National Natural Science Foundation of China (60775024).

6. REFERENCES


Figure 5. Actor correlation graphs interfaces in “Friends”. (a) and (d). A concurrence graph in all video database; (b). Correlations graph centered on Chandler; (c). Ranking shots list with a correlation query between Monica and Chandler; (e). Correlations graph centered on Rachel; (f). Ranking shots list with a correlation query between Rose and Rachel.