Synchronization among Groups of Spectators for Highlight Detection in Movies

MUSZYNSKI, Michal, et al.

Abstract
Detection of emotional and aesthetic highlights is a challenge for the affective understanding of movies. Our assumption is that synchronized spectators’ physiological and behavioral reactions occur during these highlights. We propose to employ the periodicity score to capture synchronization among groups of spectators’ signals. To uncover the periodicity score’s capabilities, we compare it with baseline synchronization measures, such as the nonlinear interdependence and the windowed mutual information. The results show that the periodicity score and the pairwise synchronization measures are able to capture different properties of spectators’ synchronization, and they indicate the presence of some types of emotional and aesthetic highlights in a movie based on spectators’ electro-dermal and acceleration signals.

Reference

Available at:
http://archive-ouverte.unige.ch/unige:85616

Disclaimer: layout of this document may differ from the published version.
Synchronization among Groups of Spectators for Highlight Detection in Movies

Michal Muszynski
Computer Vision and Multimedia Laboratory
University of Geneva
Switzerland
michal.muszynski@unige.ch

Theodoros Kostoulas
Computer Vision and Multimedia Laboratory & Swiss Center for Affective Sciences
University of Geneva
Switzerland
theodoros.kostoulas@unige.ch

Patrizia Lombardo
Department of Modern French & Swiss Center for Affective Sciences
University of Geneva
Switzerland
patrizia.lombardo@unige.ch

Thierry Pun
Computer Vision and Multimedia Laboratory & Swiss Center for Affective Sciences
University of Geneva
Switzerland
thierry.pun@unige.ch

Guillaume Chanel
Swiss Center for Affective Sciences & Computer Vision and Multimedia Laboratory
University of Geneva
Switzerland
guillaume.chanel@unige.ch

ABSTRACT
Detection of emotional and aesthetic highlights is a challenge for the affective understanding of movies. Our assumption is that synchronized spectators’ physiological and behavioral reactions occur during these highlights. We propose to employ the periodicity score to capture synchronization among groups of spectators’ signals. To uncover the periodicity score’s capabilities, we compare it with baseline synchronization measures, such as the nonlinear interdependence and the windowed mutual information. The results show that the periodicity score and the pairwise synchronization measures are able to capture different properties of spectators’ synchronization, and they indicate the presence of some types of emotional and aesthetic highlights in a movie based on spectators’ electro-dermal and acceleration signals.

Categories and Subject Descriptors
I.5.4 [Pattern Recognition]: Applications—signal processing; I.5.2 [Pattern Recognition]: Design Methodology—pattern analysis

Keywords
Synchronization; Grassmann manifolds; Persistent homology; Physiological and behavioral signals; Affective computing; Highlight detection

1. INTRODUCTION
Many studies have focused on matching spectators’ physiological signals with affective states and the appearance of highlights in movies [5]. Since physiological reactions are considered to be an important component of emotions [15], [4], their measurements provide insight into spectators’ aesthetic experience elicited by particular scenes [24]. In the field of affective computing, researchers have attempted to investigate emotion recognition in responses to multimedia content using electroencephalography (EEG) signals, peripheral physiological signals and facial expressions [14], [22]. The combination of spectators’ physiological signals has been proposed in [5]. Because the spectators were watching separately a movie without any social context, they could not interact among themselves as it is the case in our studies.

Spectators can display similar behaviors or have similar physiological reactions when they are watching a movie together because: (i) the aesthetic choices of the filmmaker are made to elicit specific emotional reactions (e.g. special effects, empathy and compassion toward a character, etc.) and (ii) watching a movie together causes spectators’ affective reactions to be synchronized through processes of emotional contagion [10]. For these reasons we gain insight on the impact of synchronization among group of spectators.

The first concept of synchronization came from the rhythm adjustment of oscillating objects [19]. In social sciences interpersonal synchrony consists of three components: rhythm, simultaneous movement and smooth meshing of interactions [2], [7]. One important step was made by [21] to introduce generalized synchronization of coupled chaotic systems. In [17], [18] the authors proposed that a level of the periodicity score can measure the amount of pattern repetitions in signals.

In this paper we propose to use periodicity score and baseline pairwise synchronization measures, such as the nonlinear interdependence [20] and the windowed mutual information [13] to uncover relations between occurrence of emo-
tional and aesthetic highlights in films and spectators’ physiological and behavioral reactions. Our goal is to verify if a level of synchronization that is computed over time windows of spectators’ electro-dermal and acceleration signals indicates the occurrence of different types of emotional and aesthetic highlights in movies. We then evaluate our results with respect to the annotations made by a movie critic.

In section 2 we detail the adaptation of synchronization measures to process spectators’ physiological and behavioral signals. In section 3 we describe a movie watching in a ecological situation. In section 4 we present all the results. In section 5 we discuss and interpret the obtained results. In section 6 we provide the conclusions of our studies.

2. SYNCHRONIZATION MEASURES

In this section we propose to apply three synchronization measures: the periodicity score, the nonlinear interdependence and the windowed mutual information to spectators’ physiological and behavioral signals in order to detect emotional and aesthetic highlights in movies. A high level of all synchronization measures reveals the synchronized reactions of spectators while watching a movie.

For spectators’ electro-dermal and acceleration signals \( \{x_i\} \) we consider time windows \( \{x_i(l)\}, \quad i = 1, ..., M, \quad l = 1, ..., N \), where \( M \) is a number of spectators’ signals and \( N \) is a number of time windows.

2.1 Contribution - Periodicity Score

In this subsection we detail the usage of the periodicity score to measure synchronization of signals \([17], [18]\) plotted in Figure 2. Once the sequence of \( S_i^t(l, f) \) matrices is mapped to \( G(k, n) \) and defines a metric space \((U = \{U^1, ..., U^M\}, d_{min}(\cdot, \cdot))\), we recall the definition of the Vietoris-Rips complex \( \text{Rips}_\alpha(U) \) as the set of the simplices \([U^1, ..., U^q]\) such that \( d_{min}(U^i, U^j) \leq \alpha \) for \( i, j = 1, ..., q \). There is an inclusion of \( \text{Rips}_\alpha(U) \) into \( \text{Rips}_\beta(U) \) for any \( \alpha \leq \beta \). The sequences of inclusions are called filtrations \( \text{Filt}_\alpha(U) \). An example is given in Figure 2(a). Persistence diagrams allow us to study the evolution of the topology of a filtration, and to capture properties of the metric which is used to generate the filtration. Existing connected components are merged for 0–th homology, when \( \alpha \) increases shown in Figure 2(b). Persistent homology tracks the birth (appearance) \( b \) and death (disappearance) \( d \) of all connected components shown in Figure 2(c).

The maximum persistence \( mp(dgm(x_i(l))) \) of a persistence diagram \( dgm(x_i(l)) \) is defined as follows \([18]\)

\[
mp(dgm(x_i(l))) = \max_{(b,d) \in dgm(x_i(l))} \text{pers}(b,d),
\]

where \( \text{pers}(b,d) = d - b \) for \((b,d) \in dgm(x_i(l))\), and as \( \infty \) otherwise. Finally, we can provide the periodicity score \( S(x_i(l)) \) \([18]\)

\[
S(x_i(l)) = \frac{mp(dgm(x_i(l)))}{\sqrt{3}}.
\]

The normalized maximum persistence \( mp(dgm(x_i(l))) \) of a persistence diagram \( dgm(x_i(l)) \) can help us to quantify synchronization among signals because it is capable of measuring their intrinsic geometric dependencies. The persistent score can measure synchronization among groups of signals based on the connectivity of signal clusters. Our approach to synchronization contains the multivariate measure which ascribes a single value to all signals in comparison with univariate measures, such as the nonlinear interdependence and the windowed mutual information which can only be computed over pairs of signals. When \( S(x_i(l)) \) equals 0, it means
that we can not explore any structure in our data. If a value of \(S(x_i(t))\) rises close to 1, we find some strong connectivity structure of data (synchronization).

![Diagram](image)

**Figure 2:** We show an example of: (a) the filtered Vietoris-Rips complex, (b) a number of connected components for different values of filtration parameter \(\alpha\), (c) the persistent diagram.

### 2.2 Nonlinear Interdependence

The nonlinear interdependence measures the geometrical similarity between the state space trajectories of two dynamical systems that are reconstructed from two time series \(\{x_i\}\) and \(\{x_j\}\) \(i,j = 1,...,M\) using time-delay embedding [23]. For each time window \(x_i(t)\) the mean square Euclidean distance to its \(K\) nearest neighbours \(x_j(n), n = 1,...,K\)

\[
R^K(x_i(l)) = \frac{1}{K} \sum_{n=1}^{K} (x_i(l) - x_i(n))^2,
\]

and the mean squared Euclidean distance conditioned by the equal time partners of the \(K\) nearest neighbours of \(x_j(l)\) is

\[
R^K(x_i(l)|x_j(l)) = \frac{1}{K} \sum_{n=1}^{K} (x_i(l) - x_j(n))^2.
\]

The nonlinear interdependence measure is defined as [20]

\[
S^K(x_i(l)|x_j(l)) = \frac{R^K(x_i(l))}{R^K(x_i(l)|x_j(l))}.
\]

To make the nonlinear interdependence symmetric, we consider \(S^K(x_j(l)|x_i(l))\) and we then average these parameters.

### 2.3 Windowed Mutual Information

From an information theory viewpoint, any signal can be treated as a collection of random variables which describes the evolution of a system over time. In this context, the windowed mutual information may capture nonlinear dependencies between signals that are not revealed in the covariance of signals [13]. The straightforward approach to estimation of the windowed mutual information consists of partitioning the supports of two time windows \(x_i(t), x_j(t)\) into finite size bins, and the approximation by the finite sum

\[
I(x_i(l), x_j(l)) = \sum_{w, q} p(w, q) \log \left( \frac{p(w, q)}{p(w)p(q)} \right),
\]

where \(p(w, q)\) is the joint probability density function of \(x_i(l)\) and \(x_j(t)\), \(p(w)\) and \(p(q)\) are the marginal probability density functions, respectively.

### 3. EXPERIMENT

The spectators’ physiological and behavioral signals were recorded with a sampling frequency of 10 Hz during a movie projection (Taxi Driver, 1976) in a theater (Gritti cinema, Geneva) [12]. The duration of the movie is 113 min. In this paper we use 12 spectators’ electro-dermal activity and acceleration \((x, y, z\) axes, accelerometer attached to the hand of the spectator) signals \((M = 12)\). All signals are filtered by third order lowpass Butterworth filter with cutoff frequency 0.3 Hz, and they are segmented into overlapping time windows with a time step and a window length equal 2 s and 5 s, respectively. The selection of these parameters was done to indicate highlights in meaningful time period for the whole duration of the movie.

Annotation of the movie was performed offline by a movie critic, who annotated the movie based on the following five types of emotional and aesthetic highlights [1], [3].

- **Form-highlights** (the manner in which the subject is presented in the film):
  - \(H1\): Spectacular (technical choice, special effects);
  - \(H2\): Subtle (use of camera, lighting, music).

- **Content-highlights** (the presented subject in the film):
  - \(H3\): Character development (characters’ emotions and responses to dramatic events);
  - \(H4\): Dialogue (motivation of actions and tensions among characters);
  - \(H5\): Theme development (unusual close up, urban theme).

In these studies the periodicity score (PS) and pairwise synchronization measures: the nonlinear interdependence (NI) and the windowed mutual information (WMI) are applied to the spectators’ physiological and behavioral signals to uncover different properties of their synchronization.

### 4. RESULTS

We employ the two side Welch’s t-test at the significance level \(\alpha = 0.1\) to test the hypothesis. We verify if an increase/a drop of the synchronization of the spectators’ physiological and behavioral reactions might appear during a particular type of highlights [16], [11].

Figure 3(a) shows the mean values of the PS which are computed over the spectators’ physiological signals for all types of highlights. The values of the PS increase marginally significantly for the scenes in the movie containing spectacular highlights \(H1\) (t=1.77, \(p<0.1, r=0.12\)) and theme development highlights \(H5\) (t=1.90, \(p<0.1, r=0.14\)) in comparison to the scenes without those particular types of highlights. Figure 3(b) plots the results for the behavioral signals. The values of the PS rise significantly and marginally significantly for the scenes consist of spectacular highlights \(H1\) (t=2.55, \(p=0.01, r=0.15\)) and subtle highlights \(H2\) (t=1.77, \(p<0.1, r=0.09\)), respectively. Moreover, the values of the PS drop marginally significantly for character development highlights \(H3\) (t=1.71, \(p<0.1, r=0.11\)).

In Figure 3(c) the mean values of the NI for the electro-dermal signals are plotted. The synchronization increases...
These observations are in line with our previous works [16], [11], [12]. This can be justified by the nature of the scenes since $H1$ corresponds to spectacular scenes where the director uses special effects, such as an increasing saturation of red color during final shooting scenes, playing with lights and a location of the camera. These may evoke strong emotional reactions and emotional contagion.

The slow movement of a camera and music during subtle highlights $H2$ could decrease the spectators’ behavioral reactions. Due to discrepancies among the spectators we observe a drop of the NI and the WMI while the PS might still capture some synchronization in their behaviors.

The PS falls marginally significantly for the spectators’ behavioral reactions during character development highlights $H5$ while the pairwise synchronization measures could not expose any marginally significant drop/increase of the synchronization. Some group of the spectators could only react similarly to the attitude of the main characters which could be caused by the ambiguity of their personality.

Also, we observe a significant drop of the WMI values for the spectators’ physiological signals in the case of dialogue highlights $H4$. This can occur because of two reasons: long average duration of highlights $H4$ may cause that the spectators’ emotions fade in time, and the main character is also an ambiguous movie character who could elicit different reactions across the audience.

The results also report that the values of the PS and NI increase marginally significantly for the theme development highlights $H5$. The rise of spectators’ synchronization might be caused partially by the overlapping of spectacular and theme development scenes in this particular movie. But the WMI also measures a significant drop in the statistical dependencies. It can be explained by a lack of any mutual dependence between signals.

5. DISCUSSION

We can detect some types of highlights applying synchronization measures. The obtained results show that the values of the PS and NI increase marginally significantly and significantly, respectively, for spectacular highlights $H1$.
7. REFERENCES