

## 視聴者の表情を用いた動画ダイジェスト作成システム

加藤 宏和 菰田 昇平 古川 裕士 森本 光紀 濱川 礼

中京大学 情報理工学部 情報システム工学科

### 1. 概要

本論文では、ユーザの表情を撮影した動画(以下、表情動画)を用いて、動画の意味やユーザの嗜好をダイジェストに反映させるシステムについて述べる。ここで、ダイジェストとは動画コンテンツの要約を指す。

本システムでは、視聴中の顔を撮影し、その表情動画を用いることで、どの部分でユーザが盛り上がったかを判定し、ダイジェストに反映するシステムを実装し、検証を行った。

### 2. 背景・目的

現在、世の中にある多くのダイジェスト作成手法は、動画中の対象物や動画の音声を解析することで盛り上がっている部分を判定している。具体的な例として、「サッカー映像のシーン自動解析の研究」[1]という研究が挙げられる。この研究は、ゴールポストが動画中に存在している部分をゴールシーンと位置づけ、ダイジェストを作成する。

しかし、これらの手法では、「意味的に盛り上がる部分は考慮されていないこと」や「多くのユーザが盛り上がるだろうという推測で判定されていること」などの問題点がある。そこで、動画視聴中の顔を撮影し、その表情動画を用いることで、ユーザの表情が変化した部分、つまりユーザが盛り上がった部分を判定し、ダイジェストに反映するシステムを開発した。

### 3. 提案手法

#### 3.1. 概要

本システムは、1台のPCと1台のWebカメラを使用する。Webカメラをユーザの顔が映るようPCに取り付ける。

図1にダイジェスト作成の利用形態を示す。右上のグラフは、赤線が笑いのダイジェストパラメータで、黒線が移動平均線を示す。本システムのプレイヤー部分で動画を再生し、同時にWebカメラでの撮影が開始される。

撮影した顔動画を解析し、表情が大きく変化した部分をデータとして作成する。

そのデータを用いて、ダイジェストを自動作成する。



図1 ダイジェスト作成の利用形態

#### 3.2. 内部処理

以降、本システムの内部処理について述べる。図2に実装したシステムでの処理手順を示す。

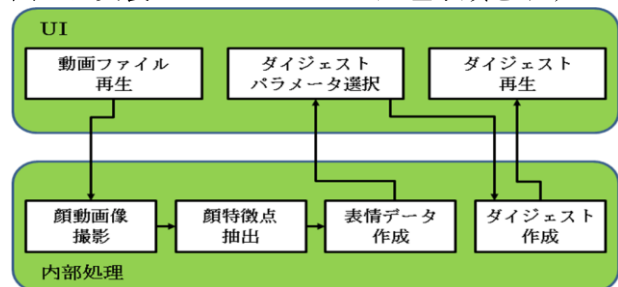


図2 処理手順

##### 3.2.1. 顔動画撮影

動画を視聴している顔を撮影する。再生終了と同時に、撮影も終了する。

##### 3.2.2. 顔特徴点抽出部

顔特徴点抽出部では、画像処理を用いて眉・目・口から表情認識に用いる特徴点を抽出する。顔画像内での顔部位の位置関係は、人やカメラ位置によってほとんど変わらない。そのため、予め決定した領域を使用して、眉、目、口の部位がある領域を切り出す。毎秒5フレームの顔動画を使用し、OpenCV[2]の学習データを用いて、顔領域を切り出す。

まず、眉領域から輝度値の小さい部分を抽出し、2値化画像を作成し、横幅の最も大きな領域から左右の端の点を取得する。次に、目領域から彩度の小さい部分(白目部分)とエッジ抽出した輪郭の2値化画像を作成し、上下左右の端の点を取得する。そして、口領域から輝度値の小さい部分と唇の色を抽出し、2値化画像を作成

The video digest creation system that uses viewer's expression ~Iitokodori(Nice scene recorder) ~  
Kato Hirokazu, Komoda shohei,  
Furukawa Hiroshi, Morimoto Mitsuki and Rei Hamakawa

し、上下左右の端の点を取得する。

### 3.2.3. 表情データ作成部

顔特徴点抽出部で抽出された特徴点座標の平均を算出する。算出された座標をユーザの基準の表情とし、各フレームとの変化量を求め、表情量として定義する。

表情量の算出は、ユーザの顔の大きさや顔の傾きなどによる表情量への影響を低減するために「顔特徴点を用いた特徴選択と特徴抽出による表情認識に基づく映像中の表情表出シーン検出」[3]で用いられている2つの方法を用いる。

- ▶ 直線距離－基準の座標の2点間の距離と各フレームの2点間の距離の差を求める。
- ▶ 三角形面積－基準の座標の3点を結んだ三角形の面積と各フレームの3点間の距離の差を求める。

盛り上がり判定する表情は、Ekmanらによって提案された基本6感情[4]の内、ダイジェスト作成に必要な感情を除いた笑い・興奮・恐怖の3種類とし、表情量の算出はそれぞれに対応した特徴点を選出する。例えば、笑いは口角が上がり目尻が下がるため、口の特徴点と目の特徴点の距離を表情量の算出に用いる。また、興奮は顔全体が外側に広がるので、眉の特徴点と口の特徴点の三角形面積を表情量の算出に用いる。図3に算出する三角形面積を示す。

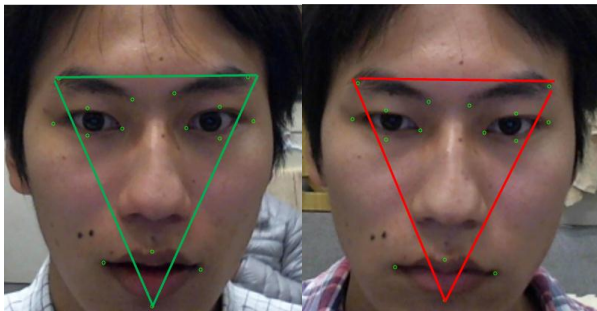


図3 算出する三角形面積

フレーム毎に算出した表情量をダイジェストパラメータに表示する。ダイジェストパラメータは、横軸に時間、縦軸に表情量として構成される。図4に笑いのダイジェストパラメータの一例を示す。

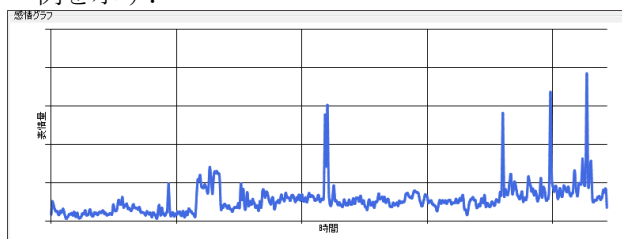


図4 ダイジェストパラメータ例

### 3.2.4. ダイジェストパラメータ選択部

ユーザに3種類から1つのダイジェストパラメータを選択してもらう。

### 3.2.5. ダイジェスト作成部

表情データ作成部で、選択されたダイジェストパラメータを用いて、ダイジェストを作成すまず、ダイジェストパラメータの値から50フレーム毎に移動平均値を算出する。次に、移動平均線よりも表情量が上回った場合、ユーザが盛り上がりしていると判断する。表情量が移動平均線を上回る点から下回った点を再生開始時間、下回った点を再生終了時間とし、繰り返し再生を行い、ダイジェストを作成する。

## 4. 評価

研究室学生24名を対象に評価を行った。15分の動画を視聴してもらい、3分ほどのダイジェストを笑い・興奮・恐怖の3種類作成し、アンケートを実施した。

3種類のダイジェスト全て、「もう一度視たいシーンがダイジェストに反映されている」という問いに対して、6割以上がそう思うと答えた。

特に笑いのダイジェストについては、感情が変化した部分を概ねダイジェストに反映させることができた。その反面、半数以上がシーンの切り替わりに違和感があると回答した。

## 5. 考察

上記の評価より、視聴者の盛り上がりを反映したダイジェストを作成できた。そして、一番重視していた映像の特徴や音声に着目する手法では取得できない部分もダイジェストに反映されていた。

しかし、動画内容を一切考慮しないため、シーンの切り替わりに違和感が生まれた。そこで、今後の展望として、違和感が生まれないシーンの切り取り方法の導入が挙げられる。

## 6. 参考文献

- [1] 山本大樹, 清水大輔他「サッカー映像のシーン自動解析の研究」電子情報通信学会究報告 2005 vol.104 no.537
- [2] OpenCV  
<http://opencv.jp/>
- [3] 野宮浩揮他「顔特徴点を用いた特徴選択と特徴抽出による表情認識に基づく映像中の表情表出シーン検出」DEIM Forum 2011C1-5
- [4] 「表情分析入門」P. エクマン他 (1987) 誠信書房出版