

facekit.net: Web から使えるカスタマイズ可能な顔画像認識器

松坂 要佐[†] 浅野 太[†]

facekit.net: A Customizable Face Recognizer on the Web

YOSUKE MATSUSAKA[†] and FUTOSHI ASANO[†]

1. はじめに

近年、非常に多くのマルチメディア情報がインターネット上で流通している。また、それらマルチメディア情報のユーザ側からのアップロードを容易にするための仕組みも盛んに開発されており、それらは広く一般化しつつある。ここではそれら Web 上で流通している文字だけではないメディアを指して Web マルチメディアと呼ぶ。

Web マルチメディアの扱いに必要とされる技術として、これまで動画の圧縮・伸張やストリーミングなどの通信技術、メディア理解に基づくインデックス化などの検索技術などが検討されてきた。我々はそれに加えて Web サービスとユーザの間のインタラクションの部分のマルチメディア化が今後重要になってくると考えている。

本研究では、それらインタラクション実現のための技術の一つとしてユーザのさまざまな情報を取得することができる顔の認識に着目し、標準的なブラウザを通してユーザの用途にカスタマイズした顔画像認識器を作成・利用することができる Web サービス facekit.net を開発した。以下その概要と応用事例を紹介する。

2. サービスの概要

サービスの概要を図 1 に示す。サービスは、顔認識機能を提供する組み込み型の認識エンジンと、ユーザにカスタマイズのための方法を提供する Web インターフェースからなる。以下それぞれの部分について説明する。

2.1 カスタマイズ用インターフェース

カスタマイズ用の Web インターフェースは、以下の部分からなる。

アカウント管理部：

[†] 産業技術総合研究所

National Institute of Advanced Industrial Science and Technology (AIST)

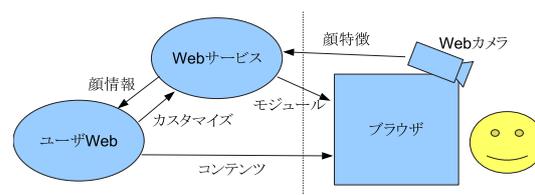


図 1 System overview

ユーザは各自のアカウントを持ちパスワードによってログインする。各々がラベル付けの作業をするプロジェクトを作成でき、他のユーザをそのプロジェクト上での作業に招待することもできる。

データアップロード部：

データアップロード部においては、動画もしくは静止画ファイルによって認識器のカスタマイズのための教示データをアップロードすることができる。データはアップロードが済んだ直後に顔検出器にかけられ、抽出された顔部分がデータベースに追加される。

ラベル付けインターフェース部：

ラベル付けインターフェース画面を図 2 に示す。ラベル付けインターフェースは、Web ブラウザ上に JavaScript によって実装されており、ユーザはドラッグアンドドロップによる直感的な操作によってラベル付け作業を進めることができる。データベースからの顔情報の取得と更新は AJAX によって実装されており、画面の更新による作業の遅延は最小で済ませられる。

ユーザは各人の用途に応じて、正面顔・非正面顔、男性・女性など任意のラベルを定義することができる。ラベル付けには名前空間も導入しており、同じ顔に対して正面 + 男性などの複数のラベルを付与することも可能である。

機械学習部：

機械学習部においてはラベル付けの終わった顔データ

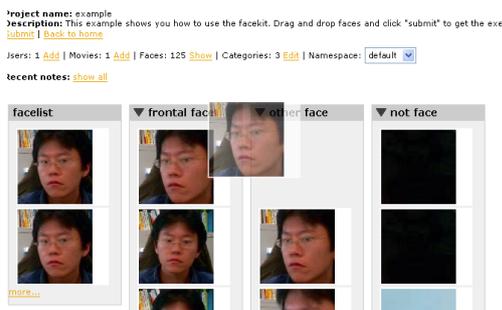


図 2 The annotaion interface

を入力として機械学習のアルゴリズムによってそれを識別するモデルを学習する。現在 SVM を用いた学習アルゴリズムが実装されており、ユーザのリクエストによって学習プロセスが開始される。この部分においては多くの計算コストが必要とされるため、ジョブ・キューを設定すると同時に、Web サーバとは別の複数のサーバにプロセスが分散される実装となっている。ユーザは E メールによって学習終了の通知を受ける。

認識エンジン配布・提供部：

学習部において作成された認識モデルは、この部分において配布・提供される。ユーザは、認識エンジンとパッケージされた認識モデルをダウンロードするか、facekit.net から提供される webAPI を経由して認識エンジンを利用するのかの、どちらかの方法を選択することができる。

2.2 認識エンジンの組み込み

パッケージ版においては、認識エンジンはモデル情報を含んだ単一の実行ファイルとして提供され、ユーザは実行ファイルの標準出力をパイプすることで各自のシステムに組み込むことができる。

webAPI 版においては、ユーザには、カメラからの映像のキャプチャと簡単な前処理を行う flash プログラムが提供され、認識はその flash プログラムと facekit.net が提供する認識サービスとの間の連携によって行われる。ユーザは各自の作る Web ページに上記 flash を貼り付けた上で JavaScript などを用いて facekit.net の指定された URL にアクセスすることによって、現在そのページをブラウズしているユーザの顔情報を得ることができる。

3. 応 用

facekit.net を用いた応用の例として、以下の二つのシステムを作成した。

・RSS リーダ

facekit のパッケージ版を用いて顔向き認識機能付きの RSS リーダを作成した。このアプリケーションにおいては、顔認識のカテゴリとして特定方向の顔とそれ以外を設定することで作成した顔向き認識器を用いてユーザの

顔向きを監視しつつ、SAPI を用いた音声合成機能によって RSS のヘッドラインを読み上げる。その読み上げに反応して顔を上げたユーザには興味があるとみなして詳細情報を読み上げる。

・広告提示

facekit の webAPI 版を用いて性別認識機能を組み込んだ広告付きの Web ページを作成した。このアプリケーションにおいては、Web ページに組み込まれた flash によってユーザの顔属性の判別を行い、その結果に応じて JavaScript によって提示される広告の内容を切り替える。

上記はどちらも、従来のマウスによる Web ページとのインタラクシオンを超えた映像メディアの入力によるインタラクシオンを標準的なブラウザの上で実現している。ユーザは facekit.net の提供するインターフェースを用いて上記の機能を容易に実現することができた。

4. 関連研究

ユーザの Web ブラウズ中の行動をマルチメディア情報を用いて評価しようという試みにはすでにいくつかの実用化の事例がある（例えば¹⁾）。

Web 上で流通するマルチメディア情報を標準的なブラウザの提供するインターフェース上で扱おうというものには、画像のラベル付けと検索機能を提供した riyu.com や like.com、ラベル付けに特化した機能を提供した labelme²⁾ や Google Image Labeler³⁾ などがある。音声認識の分野においては、近年、緒方らにより Podcast を認識・検索・訂正できる Web サービスも開発されている⁴⁾。

本研究は顔認識を行う Web サービスの構築において、1) 認識機能をリアルタイムで用いることができる API を標準的なブラウザに組み込み可能な形で提供したこと、2) ユーザによる自由なカスタマイズを可能にするラベリングインターフェースをセットで提供したこと、の 2 点において独自性がある。

5. ま と め

顔画像認識を Web 経由で作成・利用できるサービス facekit.net を開発し、いくつかの応用例を示した。

参 考 文 献

- 1) Etre - Eye Tracking, <http://www.etre.com/usability/eyetracking/>
- 2) MIT - labelme, <http://labelme.csail.mit.edu/>
- 3) Google - Image Labeler, <http://images.google.com/imagelabeler/>
- 4) 緒方ら, “PodCastle: ポッドキャストをテキストで検索, 閲覧, 編集できるソーシャルアノテーションシステム,” WISS2006 論文集 (2006).