

ARToolKit を用いた感性情報提示装置の実現

橋本 恭太

指導教員 白濱 成希

最近のネットワーク拡大によりマルチメディアコンテンツが増加している。これに伴ない現実と仮想を融合する技術の拡張現実感や、コンピューターに擬似的な感情を持たせる感性情報処理といった新しい分野の研究が行われている。これらの異なるテクノロジーを融合した新しいインターフェイスを提案する。

本稿では、ネットワーク・拡張現実感・感性情報処理のうち拡張現実感と感性情報処理に注目した。「感性君」と呼ばれる感情を表現する主観性ロボットを拡張現実感上で擬似的に表現することを目的に、感情表現の方法として色のイメージによって反応をするというシステムを構築する。

結果としては比較的人に感覚にあったものができたといえるが、今回用いた手法においては安定性など動作面において課題が残ることとなった。

1. はじめに

近年のインターネットの爆発的な普及に伴い暮らしの中でネットワーク接続は一般的なものになった。また、今までは文字情報のみで通信することが多かったがマルチメディアコンテンツ通信が多くなってきている。

その流れに、現実世界とコンピューターの生み出した仮想世界を融合する「拡張現実感」という技術がある。この技術は最近活発に行われており、ネットワーク技術と組み合わせた iPhone 用アプリ「セカイカメラ」があり注目されている。

「セカイアプリ」は、全世界のユーザーが登録したエアタグというものを iPhone のカメラ画像上で表示するというものである。これの登場によりさらに拡張現実感に関する技術は脚光を浴びている。

この拡張現実感という技術を用いることで、情報提供・作業支援・技術提供などを目的とした新しいインターフェイスを構築することが可能である。例えば、道案内表示システムや仮想的なインテリアを配置する模様替えのシミュレーターと言ったものがある。さらには、子供の情操教育として飛び出す絵本などのエンターテインメントと言った分野への発展が期待されている。

そこで本研究は感性情報提示という点に着目した。感性情報処理はロボットの発展と共に注目されていた分野であり、人間の抱く「感性」をコンピューターに擬似的に持たせ、感情を表現する。これにより親しみやすいインターフェイスが構築出来るのではないかと考えられる。

また、コストをあまりかけずに必要環境を整えることができ、個人でも比較的容易に扱える ARToolKit を用いることで幅広い層に向けて発信することが出来る。これらより、感性情報提示の一例として ARToolKit を用いた感性情報提示装置を提案する。

2. 拡張現実感 (AR) とは

拡張現実感とは Augmented Reality という物であり、現実からの知覚情報にコンピューターなどの電子計算機が作り出した電子情報を重ねる事によって現実世界の情報を強化するという技術である。この特徴から強化現実とも呼ばれている。例を上げれば現実情報の地図に電子計算結果である道のりを付加するというカーナビも AR 技術の一つであると言える。

似た技術に仮想現実 (Virtual Reality: VR) と呼ばれる技術があるが、これと異なる点がある。それは現実世界には存在していない物がそこに存在しているような体験をすることが出来る点である。また仮想現実空間すべてが電子情報であるのに対し、拡張現実感には現実世界の上に成り立っている。

現在 AR 技術の主流は視覚情報を拡張する研究である。この AR 技術にはマーカーを用いてマーカー認識によって表現する ARToolKit とマーカーを用いず空間認識によって表現する PTAM がある。ARToolKit は動作が軽く安定しているという利点があるが、マーカーを用いるので場合によっては AR らしさが失われるという短所がある。PTAM については、空間認識によって AR を実現しているためより AR らしさを表現出来るという利点があるが、ARToolKit に比べ動作が重く不安定である。

また、AR 技術の研究は視覚以外に聴覚・嗅覚・味覚・触覚の五感を拡張することを目的にしている。今後の研究でそれら五感すべてを拡張出来る時代が来ると期待されている。

3. システム概要

開発環境は WindowsXP SP3 で VisualStudio 2008 を用いて C++ で開発した。

本研究におけるシステムは色識別を OpenCV で行い、感情表現は ARToolKit を用いて AR 上で表現する。

また、このシステムは同研究室で研究していた「感性君」をモデルとしている。「感性君」とは

主観性エンジンを搭載して擬似人格をもたせた主観性ロボットのことである。

3.1. ARToolKit

3.1.1. ARToolKit とは

本研究では ARToolKit と呼ばれるプログラミングライブラリを用いる。これは AR アプリケーションの実装を手助けするライブラリであり、Windows や Linux や Mac などプラットフォームに依存しない。このライブラリは奈良先端科学技術大学院大学の加藤一博教授によって開発された。

ARToolKit は次のような機能より AR を表現する。

- ・ カメラからの画像の取得
- ・ マーカーの検出とパターン認識
- ・ マーカーの 3 次元的位置・姿勢の計測
- ・ 実写画像に 3DCG をオーバレイ表示

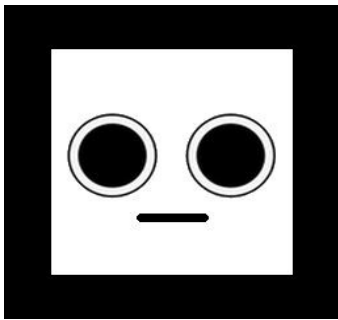


図1:作成したマーカー

ARToolKit に用いるマーカーは図 1 に示す。図 1 は一例であり、マーカーはユーザーが任意に作成することが可能である。しかし、制約がありマーカーは正方形でパターンは太枠の中に重ならないようにしなければならない。また複雑なデザインは検出精度が落ちてしまう。点対称なものも認識時の座標系の値が反転してしまい安定しなくなってしまうが、パターン認識は行われる。

次に ARToolKit 内部での処理のうちマーカーの検出時に行う 3 次元位置・姿勢の計測について説明する。

3.1.2. マーカーの検出

マーカーを検出する際にはいくつかの処理を経る。まず、入力画像を 2 値化しラベリング処理を施しラベルを割り振っていく。この領域ごとに端の点を探し、そこを起点とし領域の輪郭抽出を行う。これにより 4 つの頂点を持つ輪郭を抽出し四角形と認識する。最後にこの四角形の中に書かれているパターンをテンプレートマッチング法と呼ばれる方法で認識している。

3.1.3. マーカーの計測

ARToolKit の処理の中で核となっている処理がマーカーの位置・姿勢の計測である。検出したマーカーからマーカー・カメラ間の座標変換行列を求める。カメラ座標系を $[X_c \ Y_c \ Z_c \ 1]^T$ マーカー座標系を $[X_m \ Y_m \ Z_m \ 1]^T$ と置き、マーカーとカメラ間の座標変換式は(1)式のように定義する。

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_m \\ Y_m \\ Z_m \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{3 \times 3} & \mathbf{T}_{3 \times 1} \\ 000 & 1 \end{bmatrix} \begin{bmatrix} X_m \\ Y_m \\ Z_m \\ 1 \end{bmatrix} = \mathbf{T}_{cm} \begin{bmatrix} X_m \\ Y_m \\ Z_m \\ 1 \end{bmatrix} \quad (1)$$

計算の流れは、座標変換行列 \mathbf{T}_{cm} を求めるのに回転行列 $\mathbf{R}_{3 \times 3}$ と並進ベクトル $\mathbf{T}_{3 \times 1}$ が必要となるのでこれらを求める。

まず始めにマーカーの姿勢を表す回転行列を求める。マーカーの平行な辺の組より、画像平面に映ったマーカーの辺とカメラの焦点を通る面の式を計算し、面の法線ベクトル $\mathbf{V}_1, \mathbf{V}_2$ を求める。その後 $\mathbf{V}_1, \mathbf{V}_2$ に直交する方向ベクトル \mathbf{V}_3 を求める。これによりマーカーの姿勢を表す 3 つの方向ベクトルが求まり、これらのベクトルを並べたものが回転行列となる。

$$\mathbf{R}_{3 \times 3} = [\mathbf{V}_1 \ \mathbf{V}_2 \ \mathbf{V}_3] \quad (2)$$

次に、マーカーの座標を表す並進ベクトルを求める。マーカーをカメラ座標系の原点におき、この状態から(2)式で求めた $\mathbf{R}_{3 \times 3}$ によって回転させ、マーカーの 1 辺の長さともマーカーの頂点の 3 次元座標を用いて並進ベクトル $\mathbf{T}_{3 \times 1}$ を求める。

$$\mathbf{T}_{3 \times 1} = [t_x \ t_y \ t_z]^T \quad (3)$$

これらの計算を何度も繰り返すことによって誤差をなくし、マーカー・カメラ間の座標変換が行われマーカーの姿勢・位置を検出することができるのである。

これらの処理を行うことでマーカーを基準としたカメラの位置・姿勢も得ることができ、現実のカメラの動きに連動した仮想世界の表現が可能となる。これにより 3D オブジェクトの見え方が現実世界と同じものに見えるようになる。

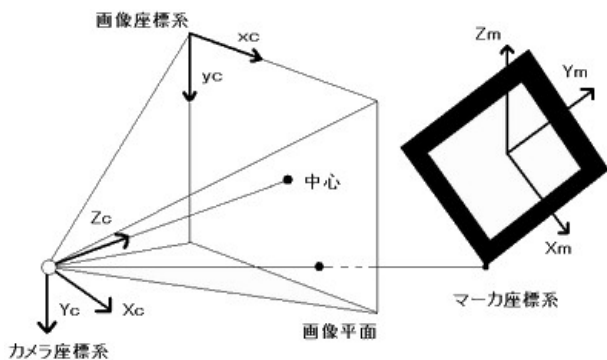


図2:カメラ・マーカ座標系

3.2. OpenGL とは

OpenGL とは Silicon Graphics 社が中心となって開発したプログラミングインターフェイスである。これは 3DCG の標準グラフィックスライブラリであり、CAD システムなどに使用されている。これは様々なプラットフォームに対応しており、簡潔なコマンド体系をとっており、高度な 3DCG 描写能力があるという特徴を持っている。

ARToolKit の 3DCG をオーバーレイ表示する際にもこの OpenGL が用いられている。ARToolKit で計算された、マーカ・カメラ間の座標変換によって求められたカメラ座標を用いて 3DCG の描写を行う。

3.3. OpenCV

3.3.1. OpenCV とは

OpenCV とはインテルが開発・公開したオープンソースのライブラリである。このライブラリは画像処理や画像認識に向いており、デジタルカメラの顔認識などによく使われている。

本研究では、色識別をこの合うために OpenCV を用いた。色抽出のために背景差分法を行った。

3.3.2. 背景差分法

背景差分法とはカメラが固定されていて視界の移動がない場合に、画像中から注目すべき物体を抽出する方法である。背景の情報を記憶し、現在の画像の各画素値が背景に属しているか否かを判断し背景と前景に分割するという手法である。

RGB の各成分の絶対値を評価し、絶対値が異なる部分は前景である単純な背景差分法を用いた。しかしカメラで観測する画像は同じ場所を撮影し続けたとしても、微妙に明るさが変化したり植物のように微妙な揺れを繰り返す対象物が映る場合、それを前景とみなしてしまう。

そこで背景の揺らぎをモデル化することでこれを解消することにした。背景の揺らぎは単一の正規分布をなすモデルと仮定し背景モデルを構築した。RGB の画素値のみを評価基準とした

場合には画像全体をひとつの閾値で前景と背景の評価をしていたが、背景をモデル化することで画素毎に閾値を変化させることが出来るようになった。



図 3-1:背景画像

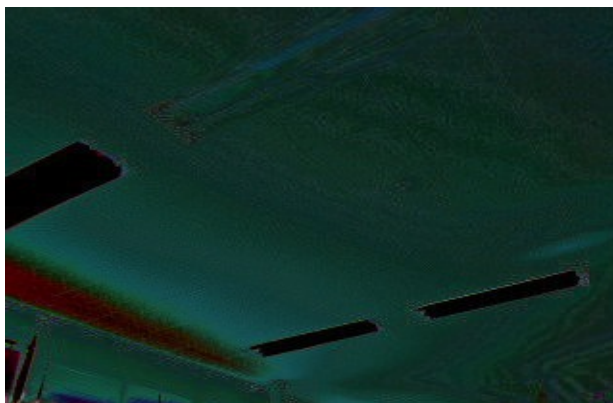


図 3-2:モデル化した背景画像



図 3-3:入力画像



図 3-4:背景モデルを考慮しない前景抽出



図 3-5:背景モデルを考慮した前景抽出

図 3-4, 3-5 に背景モデルを考慮していない場合と背景モデルを考慮した場合の違いを示す。二つの図から明らかなように背景モデルを考慮したほうがノイズがなくなっている。このノイズの原因は撮影中に太陽光の明るさが変化したことによるものである。

図 3-4 は図 3-1 と図 3-3 との差分のみで生成し、図 3-5 はそれに加え図 3-2 を考慮し、閾値を変化させて生成している。これにより、背景モデルを用いた背景差分は安定することが分かる。

4. 感情モデル

本研究における各感情を定義するのに Pultchik の感情理論を用いる。Pultchik は「日常生活の中に見られる複雑な情動は、いくつかの因子に分けられ、これはまた総合も出来る」という説を提唱した。

Pultchik は多次元尺度構成法によって 8 つの基本感情とその属する感情を立体化した。この立体モデルは横の断面には 8 つの基本感情が配置され、縦の次元では強度を表している。また次のような 6 つの公準によって作られている。

- (1) 数個の基本(1次)感情がある
- (2) 基本感情の結合の仕方によって複雑な混合感情が生まれる
- (3) 各基本感情は、その生理的行動において特異である

- (4) 日常生活で普通われわれのしている感情は混合感情であり、基本感情は混合感情から推測される
- (5) 基本感情は両極的におかれ、2つの感情がそれぞれ対になって位置付けられる
- (6) 感情はそれぞれ強度差あるいは覚醒差がある



図 4:Plutchik の感情モデル

本研究では公準 1 から基本感情として喜び・悲しみ・怒り・恐れ・期待・驚き・受容・嫌悪の計 8 つの感情を用意した。また、本研究ではモーションによる感情表現を行うが、強度ごとのモーションを設定するのは難しいと判断し公準 6 は無視した。

5. AR based KANSEI Interface の開発

感情を表現するには各感情の 3D モデルのモーションが必要になってくる。これはモデリングソフト「Metasequoia」でモデルを作成し、「RokDeBone2」でモデルのモーションを作成する。

表情に関しては、各感情の最大強度のものを用意している。各感情の表情は表 1 の通りである。また、モーションも各感情の強度を最大として作成する。

表 1:表情のイメージ

喜び	悲しみ	怒り	恐れ
期待	驚き	受容	嫌悪

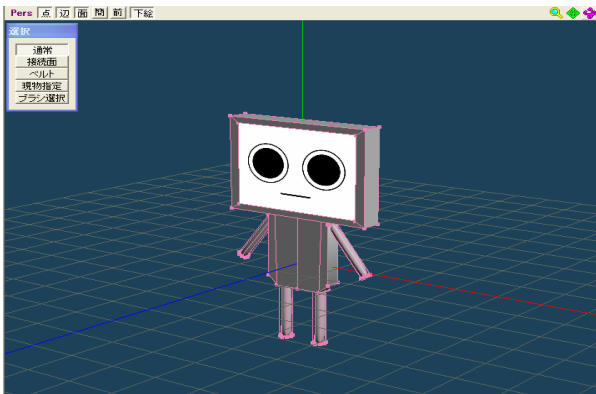


図 5:メタセコイアでのモデル作成画面

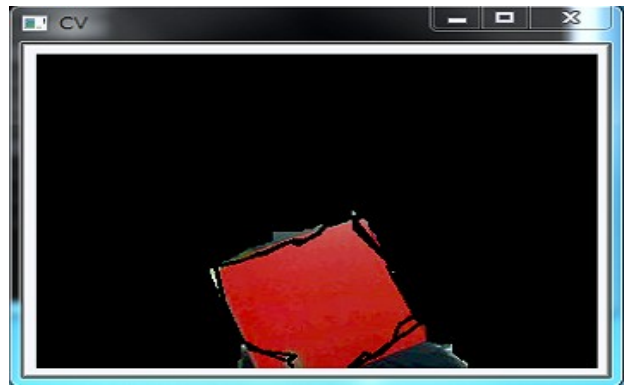


図 6-3:抽出した前景画像

6. 実際の動作風景

プログラムを実行すると以下の処理を行う

- ・ カメラパラメータや感情モデルのロード
- ・ 背景モデルを生成するため 100 枚の画像を取得し、背景モデルを生成する。
- ・ 生成した背景モデルから、現在の USB カメラ画像の各画素値を評価し前景を抽出する
- ・ 抽出した前景画像の色相値を用いてヒストグラムを生成し、最も頻度の高い色を取得する
- ・ 取得した色から対応する感情モデルを ARToolKit により描写する

プログラムを実際に動かした時の説明をする

図 6-1 は 100 枚の画像の平均から生成された背景画像である。そして図 6-2 は現在の USB カメラの入力画像を表す。図 6-1 と図 6-2 で背景差分法を用いて前景画像を抽出したのが図 6-3 である。図 6-4 は抽出した前景画像の色相値をヒストグラム化したものである。

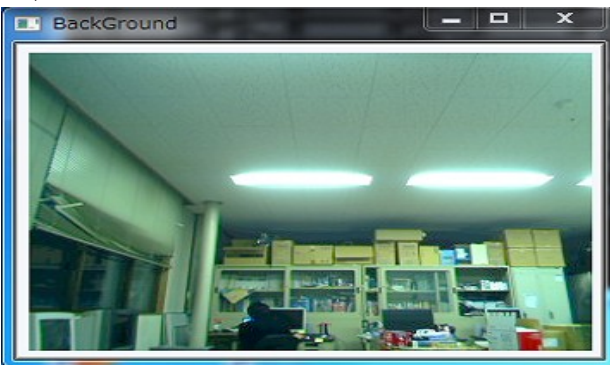


図 6-1:背景画像

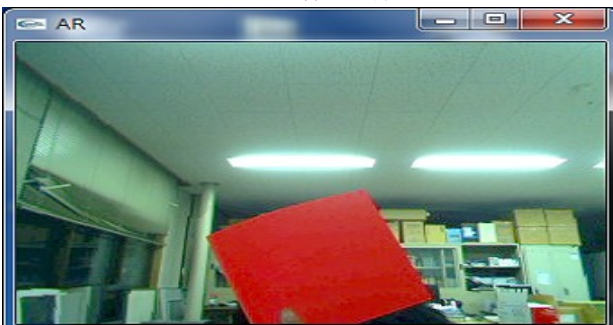


図 6-2:現在の入力画像

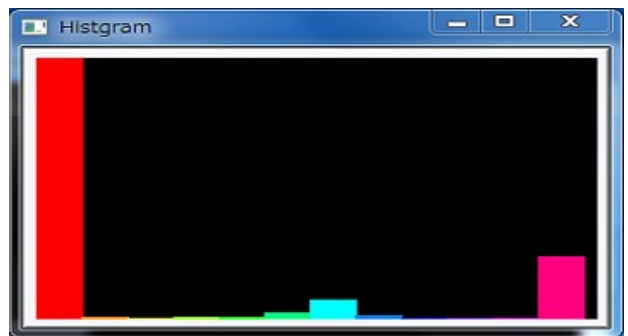


図 6-4:前景画像の色相ヒストグラム

また、色と感情の組み合わせは同研究室の森の行った WebClass でのアンケートを参考にした。その組み合わせは表 2 に示す。

表 2:色と感情の対応表

黄色	緑	青	赤	紫	黒	白	青	赤	黄	黒	
喜び	受容	受容	悲しみ	悲しみ	恐れ	恐れ	嫌悪	嫌悪	怒り	怒り	期待

図 7 は実際にプログラムを実行した画面である。この画面は、黄色の色紙に反応し、喜びの感情を表している。ヒストグラムでも黄色が最大の値をとっている。しかし、照明光や影が入ると色相値が変化してしまい、色をきちんと認識しないことがあった。

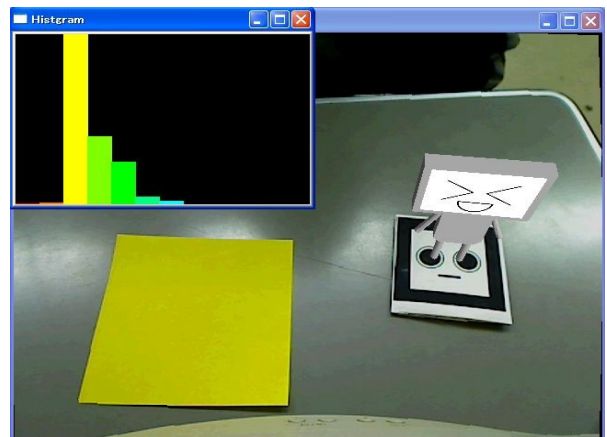


図 7:実際の動作画面

7. 考察

本研究の評価には、本科生27名に対してアンケートを実施した。アンケートの形式は被験者に、用意した色に対して抱くイメージと実際にシステムを動かしたときに挙動があっているかどうかを「あっている」「だいたいあっている」「あまりあっていない」「あっていない」の4段階で評価してもらった。

結果はあっているが5人、大体あっているが20人、あまりあっていないが1人、あっていないが0人となった。しかし、分からないという回答も1人あった。アンケートの結果からは全体の約75%がプラスの評価を得ることができた。この結果から比較的人の感覚にあったものが構築できたと考える。

しかし、モーションは色のイメージと合っていたようだが、感性は人それぞれのものなので、あっていないという意見もでた。

またモーションがあっていないと言うよりは色と感情があっていないと言う意見もでていた。

これらの意見を参考に今後の研究に活かしていきたい。

8. 今後の課題と目標

課題点を挙げる。

(1) 画素値の変化

照明や太陽などの源光の明るさが変わってしまい、カメラの自動調節によってカメラパラメータが変化してしまうという問題が発生した。

これは背景画像を随時更新することで解決出来ると考えたが、前景として注目したい部分が背景となってしまう事が考えられる。そのため、他の方法を模索しなければいけないと考える。

(2) 背景差分法の欠点

今回は前景を抽出するのに背景差分法を用いたが、これではカメラを固定しなければいけない。そのために可視範囲が縛られてしまうし、マーカーも動かすことができないという欠点がある。特に後者はARToolKitの特性上解決することは非常に難しいと考えられる。これらを解決するにはPTAMを導入することで多少は緩和されると考える。

(3) ヒストグラム化における問題

色相値をヒストグラム化しているため、少しの色でも頻度最大とみなされ抽出される。そのため、感情の強度をつけることが難しくなってしまうし、常に感情を表現し続ける事になってしまう。

これを緩和するには、前景画像の大きさがが全体画像のどれぐらいの割合にあたるかを判断材料にすればいいと考える。

(4) 感情の強度と感情の種類

本研究では感情の強度、複合感情という概念を無視している。また、色と感情に関するアンケートでは驚きに値する色が得られなかった。

これらのことより細かな感情表現は行えなかった。そこで、感情の強度や、複合感情を考慮したモーションを作成することによって、感性をよりリアルに表現できるのではないかと考える。

また、色と感情のアンケートも色の種類を増やして再度アンケートを取っていきたい。

(5) 今後の目標

今回は、色識別のみによる感情表現を行った。

また、色のみではなく対話や楽器などの音声情報もしくは、ニュースや占いといった文章のような活字情報による感情表現を行えるように目指したい。

最終的には、感性情報処理と拡張現実感にさらにネットワークを融合した技術を実現したいと考える。

本研究では、「感性君」を拡張現実感上で表現した。拡張現実感で表現する理由としては、本物のロボット「感性君」を必要としないで済むことや、プログラムによって構築されているので様々な拡張を行うことができる点がある。また、コンピューター上で動作しているので様々なシミュレーションが可能であるという点が上げられる。

9. 謝辞

最後に、本研究を行うにあたり多大なるご指導、ご鞭撻下さった白濱成希教員に厚くお礼申し上げます。また、アンケートにご協力下さった本科生の皆様にも厚くお礼申し上げます。

参考文献

- [1]橋本直, 「ARToolKit 拡張現実感プログラミング」, アスキー・メディアワークス, 2008
- [2]林武文, 加藤清敬, 「OpenGLによる3次元CGプログラミング」, コロナ社, 2003
- [3]谷尻豊寿, 「ARToolKitプログラミングテクニック」, カットシステム, 2008
- [4]奈良先端科学技術大学院大学 OpenCV プログラミングブック製作チーム, 「OpenCV プログラミングブック」, 毎日コミュニケーションズ 2007
- [5]VINTAGE, 主観性表現エンジン感性君 <http://www.vintage.ne.jp/works/kanseikun>
- [6]石井さや, 「感情表現を重視した表情インターフェイスに関する考察」, 北九州工業高等専門学校専攻科特別研究論文, 2006