

ARToolKit を用いた感性情報提示装置の実現

桑名修平

指導教員 白濱成希

ネットワークの普及につれマルチメディアコンテンツが増加している。また、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も行われている。これらのテクノロジーを融合することで新しいインターフェースを提案する。

本稿ではこれら3つの内、拡張現実感、感性情報処理に焦点を置いた、「感性君」とよばれる感情を表現する機能を搭載した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からのイメージによって反応を返すというシステムを構築する。結果としては人の感覚にあったものができたといえる。しかし、今回用いた手法では安定性などの面において課題が残る結果となった。

1. はじめに

インターネットの普及に伴い、暮らしの中でネットワーク接続が前提となる場面が増加している。また旧来の文字情報だけの通信ではなく、マルチメディアコンテンツ通信も身近なものになってきている。

その流れの一つに、現実世界とコンピュータが作り出した仮想世界とを融合する技術として拡張現実感と呼ばれている技術がある。この技術の研究が近年より活発に行われている。

最近では、iPhone 専用アプリの「セカイカメラ」が公開されて脚光を浴びた。カメラで撮影している映像内の場所や物にエアタグと呼ばれる付加情報がオーバーレイ表示されるアプリケーションである。このエアタグはユーザー自身が作成でき、コメントを書いたり、画像を貼り付けることができる。

この拡張現実感という技術を用いることで、情報提供や、作業支援、技術提供等を目的とした新しいインターフェースを構築することが可能である。例えば、緊急時の避難経路の表示や、仮想的にインテリアを配置して模様替えができたこと、医療分野では手術シミュレータといったことも実現できるであろう。さらには、子供の情操教育として飛び出す絵本の開発やエンタテインメント、といった分野への展開も期待されている。

本研究では感性情報提示という点に目をつけた。感性情報処理は以前からも注目されていた分野で、人間が抱く「感性」をコンピュータに付加する技術である。感情は対話などのやり取りにおいて、重要な役割をするので、コンピュータに擬似的な感情を持たせることでより親しみやすいインターフェースを構築することが可能になると考えられる。

ARを用いる利点として、コストをあまりかけずに必要環境をそろえられる、個人でも比較的簡単に扱える、PC上で動くのでシミュレーションが可能といった点が挙げられる。このことか

ら、本研究では感性情報提示の一例としてARToolKitを用いた感性情報提示システムを提案する。

2. AR(Augmented Reality)とは

ARとは拡張現実といわれる技術である。この技術は、現実からの知覚情報に、コンピュータを用いて作り出した電子情報を重ねることによって現実世界の情報を強化するというものである。この特徴から、強化現実とも呼ばれる。

似た技術に仮想現実(VR)と呼ばれるものがあるが、これと異なる点は現実に存在しないものがあたかもその場にあるような体験をすることができるという点である。

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技術にはマーカ型のARToolKitとマーカレス型のPTAMがある。マーカ型の利点は動作が軽く比較的安定しているという点である。それに対しマーカレス型はマーカを使用せず場所を選ばないために現実感を表現するのに向いていると言える。本研究では動作が安定しているマーカ型のARToolKitを用いる。

そもそもARとは視覚以外に聴覚、嗅覚、触覚、味覚の五感すべてを拡張することを目的としている。今後の研究で、それら五感すべてを拡張する時代が来ると期待されている。

また期待される技術として、このARをヘッドマウントディスプレイに導入することで、よりリアルな拡張現実の表現が可能だといえる。さらに、拡張現実コンタクトレンズという物の研究も進んでいる。

3. 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。このARToolKitとは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る。マルチプラットフォームでありWindows, LinuxなどさまざまなOSに対応している。これを用いることでパターンをカメラで読

み取り，マーカ上に3DCGをオーバーレイ表示するアプリケーションを作ることが出来る。

これは，奈良先端科学技術大学院大学の加藤博一教授によって開発された。

ARToolKitのおもな機能は，

- ・ カメラからの画像の取得
 - ・ マーカの検出とパターン認識
 - ・ マーカの3次元位置・姿勢の計測
 - ・ 実写画像と3DCGのオーバーレイ表示
- のような処理がある。

ARToolKitで実際に用いるマーカを図1に示す。

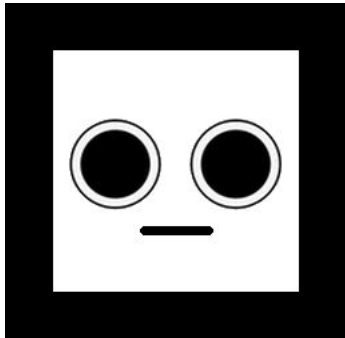


図1，作成したマーカ

このマーカはARToolKitに付属されているツールによってユーザーが任意に作成可能である。パターンは黒い太枠の中に書きさえすればよい。ただし，複雑なデザインであると検出精度が落ち，また，点对称なものは認識時の座標系の反転を招くためこれらは避けたほうがよい。

ARToolKit内部での処理のうちマーカの検出，3次元位置・姿勢の計測について説明する。

マーカを検出する際にはいくつかの処理を経てマーカの領域を求めている。まず，2値化処理を施し白黒画像に変換し，ラベリングという処理をして領域ごとにラベルを割り振っていく。この領域ごとに端の点を探し，そこからマーカの領域を抽出する。そして，最後に四角のなかのパターンをテンプレートマッチング法と呼ばれる原画像と比較し検出を行っていく手法を取りマーカを認識する。

ARToolKitの処理の中で核ともいえる処理がマーカの位置・姿勢の計測である。拡張現実感を違和感なく表現するにはカメラに対するマーカの3次元位置・姿勢を求めなければならない。そこで，arGetTrnsMat()関数を用いて，検出したマーカからマーカ・カメラ間の座標変換行列を求める。このシステムの座標系を図2に示す。

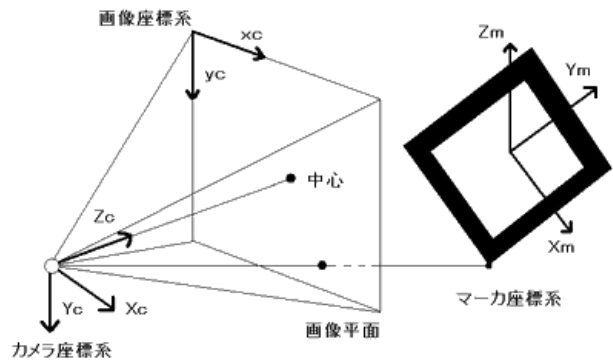


図2，システムの座標系

カメラ座標系を $[X_c \ Y_c \ Z_c \ 1]^T$ ，マーカ座標系を $[X_m \ Y_m \ Z_m \ 1]^T$ と置くとマーカとカメラ間の座標変換は(1)式のように定義できる。

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

また，画像座標系を $[x_c \ y_c \ 1]^T$ とおくとカメラ座標系から画像座標系への射影変換を(2)式のように定義できる。ここで(2)式の h はスカラーを表していて，パラメータ \mathbf{P} はカメラキャリブレーションによって求められる。

$$\mathbf{P} = \begin{bmatrix} P_{11} & P_{12} & P_{13} & 0 \\ 0 & P_{22} & P_{23} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} h_{xc} \\ h_{yc} \\ h \\ 1 \end{bmatrix} = \mathbf{P} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \quad (2)$$

実際の画像平面上での座標値は(3)式で求められる。

$$x_c = \frac{P_{11} X_c + P_{12} Y_c + P_{13} Z_c}{Z_c} \\ y_c = \frac{P_{22} Y_c + P_{23} Z_c}{Z_c} \quad (3)$$

計算の流れとしては，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 $\mathbf{T}_{3 \times 1}$ が必要になるため，まずこれらを求める。

回転行列はマーカの姿勢を表す3つの方向ベクトルからなり(4)式で表される。ここでいう(4)式は最終的な値ではない。ARToolKitではこれを計算の初期値とし，リアルタイムで計算を繰り返すことで誤差が小さくなるようにしている。

$$\mathbf{R}_{3 \times 3} = [\mathbf{V}_1 \quad \mathbf{V}_2 \quad \mathbf{V}_3] \quad (4)$$

並進ベクトルはマーカを並進させるためのもので(5)式で表される。

$$\mathbf{T}_{3 \times 1} = [t_x \quad t_y \quad t_z]^T \quad (5)$$

以上の計算により，各座標系への変換が行われ，画面上にマーカを射影することができるのである。

これらの処理を行うことで仮想カメラを，現実世界のカメラの位置・姿勢・パラメータと等しくすることができる。すなわち，現実世界のカメラの動きに連動した仮想世界の表現が可能になる。これにより，3Dオブジェクトの見え方が現実世界でのものと同じになりあたかもその場にあるような感覚が得られるのである。

以上にあげた処理を実際に行うには専門的な知識が必要である。しかし，ARToolKitはこれらの処理をブラックボックス化してくれる。そのため，たとえ座標変換という処理の理解をしていなくても拡張現実を簡単に実現できる。

4. 感情モデル

本研究において各感情を定義するのに際し，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て，これはまた総合もできる」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。そ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ており，縦の次元で強度を表している。なお，次のような6つの公準を基に作られている。

- (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

本研究では公準1の「数個の基本(1次)感情がある」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。また公準5から，基本感情のベク

トル配置は図3に示す通りとする。公準6に「感情はそれぞれ強度差あるいは覚醒差がある」とあるが，本研究ではモーションにより感情表現を行うのだが，強度ごとにモーションを設定するのは難しいと判断し，最大強度での表現のみを行うこととする。



図3. 感情の環

5. OpenGL とは

OpenGLとはSilicon Graphics社(SGI)が中心となって開発したプログラムインターフェイスである。これは，三次元CG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。OpenGLのほかにMicrosoft社の提供するDirect Xがある。様々なプラットフォームに対応している。簡潔なコマンド体系をとっており，三次元CGの高度な描写能力があるといった特徴をもっている。

ARToolKitには3DCGの操作を行う関数が含まれていない。そこで，このOpenGLを用いる。3DCGの表示，拡大・縮小や移動といった描画処理に関しては全てOpenGLまたは，その補助ライブラリであるGLUT(OpenGL Utility ToolKit)を用いる。

6. 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コンピュータビジョン向けライブラリである。このライブラリを用いると，画像処理等のプログラムが組める。

本研究では，色識別を行うためにOpenCVを用いた。背景差分法を用いて，USBカメラの映像から前景を抽出した。

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手法である。背景の情報を記録し，現在の画像の各画素が背景に属しているか否かを判断し前景と背景とに分割する手法である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という単純な背景差分法を用いていた。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，微妙に明るさや植物のように微妙な揺れをををする対象物が映

る場合それを前景とみなしてしまう。これを解消するために背景の揺らぎをモデル化した。背景の揺らぎは単一の正規分布をなすモデルと仮定して背景モデルを構築した。これにより単純な背景差分法では画像全体を一つの閾値で判断していたが、背景モデルを用いることで画素ごとに閾値を帰ることが出来るようになった。

7. システム概要

開発環境は WindowsXP SP3 で VisualStudio2008 を用いて C++ で開発した。

本研究におけるシステム概要を図 4 に示す。は色識別による感情の表現を OpenCV と ARToolKit を用いて行うものである。このシステムの大まかな動作は、OpenCV を用いて USB カメラからの入力画像を前景と背景に分離し、分離した前景画像から色を抽出し、それに応じた感情モデルを ARToolKit で拡張現実上に表現するというものである。

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」とは研究が行われている主観性エンジンと呼ばれるパラメータを設定することで簡単に擬似人格を搭載できるロボットのことである。この「感性君」には色を見てそれに対し設定した主観に応じた反応を示すという機能がある。

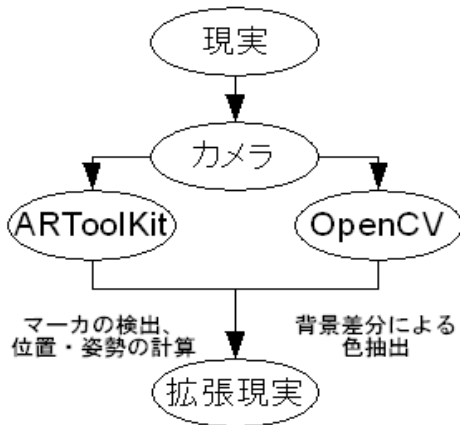










図 4, システム概要

8. AR based KANSEI Interface の開発

感情表現をする際には各感情の 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、モーション作成ツール「RokDeBone2」でモデルのモーション作成をする。

表情に関しては、各感情の最大強度のものを使用する。各感情の表情は表 1 に示す。図 5 は無表情時のモデルを表している。また、モーションも各感情の強度を最大にして作成する。

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
			

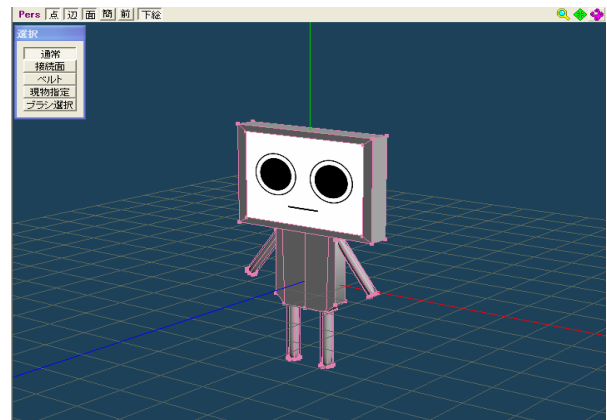


図 5, メタセコイアでのモデル作成画面

次に各モーションの挙動について説明する。

- ・ 喜び
万歳のモーション
- ・ 悲しみ
膝から崩れ落ちて泣くモーション
- ・ 怒り
モーションではなく腕を組んで仁王立ち
- ・ 恐れ
ガクガク震えるモーション
- ・ 期待
腕をばたつかせ落ち着きのない様子
- ・ 驚き
びっくりして飛び退くモーション
- ・ 受容
手を胸の前で広げて受け容れる様子
- ・ 嫌悪
首を横にふって嫌がっているモーション

9. 実際の動作

プログラムが実行されると以下の動作を行う。

- ・ カメラパラメータや、感情モデルのロード
- ・ 背景となる画像を 100 枚の画像を取得しその平均を用いることで生成する
- ・ 生成した背景画像と現在の USB カメラ画像との差分を取り、前景画像を取得する
- ・ 前景画像の色相値によってヒストグラムを生成し、もっと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

プログラムを実行した際に生成される各ウィンドウについて説明する。図6はUSBカメラからの映像を表している。このウィンドウが現実世界の映像である。図7はプログラムを実行した際に取得した100枚の画像の平均を取り生成された画像である。背景差分法を行う際の基準となる画像である。図8は前景画像である。背景差分法を行い、背景画像とカメラからの入力画像との差分をとって変化があった領域だけを抽出した画像である。図9はヒストグラムを表している。前景画像の色相値をヒストグラム化したものである。

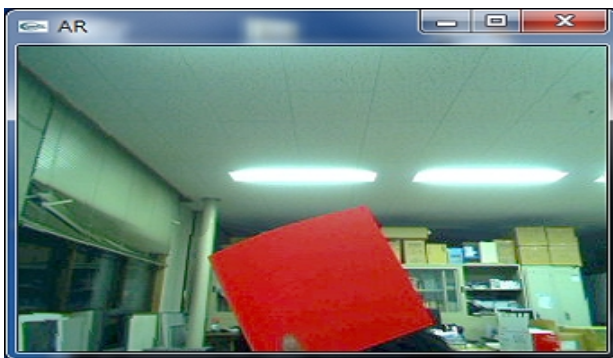


図6, カメラからの入力映像



図7, 背景画像

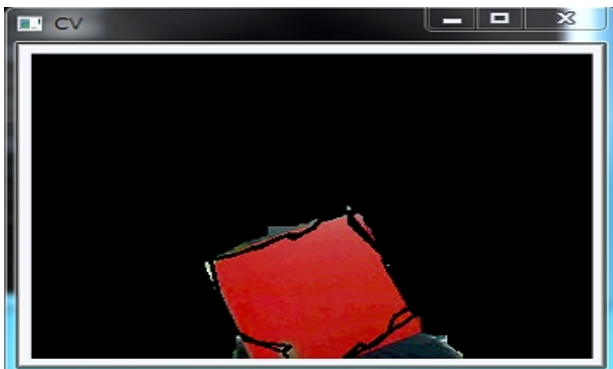


図8, 前景画像

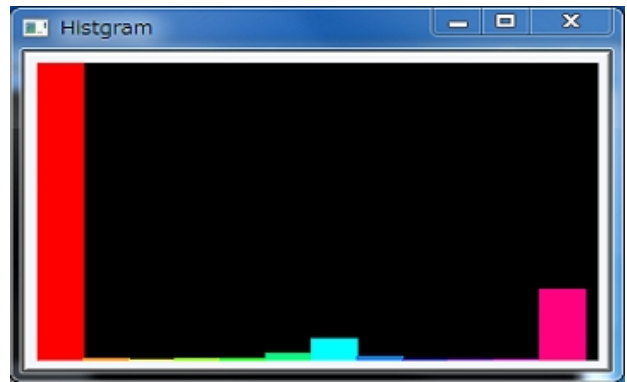


図9, 色相値のヒストグラム

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

表2, 色と感情の対応表

黄色	緑	青	水色	赤	紫	黒	白	赤	青	黄	
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10は実際にプログラムを実行した画面である。紙面上ではわかりにくいですが黄色の色紙に反応して、喜びの感情をあらわしている。ヒストグラムでも黄色が最大の値をとっているので正しく動作しているといえる。また、色紙を変えてもその色に対応した挙動を返してくれたことも確認できた。ただし、色紙に照明光や影が当たると人間の目から見ても大差なくても、PC内部では実際の色紙よりも明るい色や暗い色としてとらえてしまい正しい挙動をしない事もあった。

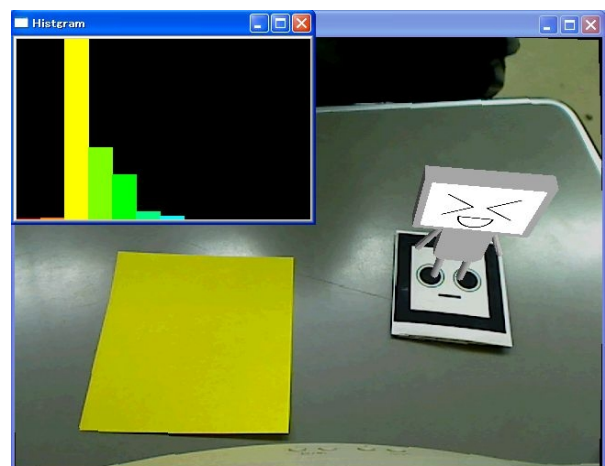


図10, 実際の動作画面

10. 考察

本研究の評価を行うに際し、本科生27人に対し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際の挙動とが合っているかどうかを調べている、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うという方式にした。

結果は、あっているが5人、だいたいあっているが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、うち1人はわからないとの回答であった。

アンケート結果からは全体の約90%の被験者からプラスの評価をいただくことができたといえる。この結果から比較的人の感覚にあったものができたと考える。

ただ、大体のモーションは色のイメージとあっていたようだが、感性はやはり人それぞれのように、あっていないという意見もでてきた。今回用意した感情のうち「受容」をモーションで表現するのが難しかったため、この感情があまり評価がよくなかったという結果もでた。

また、モーションがあっていないというよりは色と感情があっていないという意見もでてきた。さらには、色だけで感情を表現するという事がよくわからないという意見もあった。これらの意見を留意して今後の研究に活かしていきたい。

11. 課題, 目標

課題点を以下に挙げる。

(1) 画素値の変化

照明で影ができたり。カメラが明るさ調整をしてしまうと画素値が変化してしまうという問題があった。これは、背景画像をリアルタイムで生成すれば解決すると考えられるが、これではシステムの挙動がかなり重たくなってしまうことが容易に考えられる。そのため、別の方法を模索しなくてはならない。

(2) 背景差分法の欠点

今回のシステムには背景差分法を用いているためにカメラを固定していないといけない。そのために行動範囲が縛られたり、マーカすら動かすことができないという欠点が浮上してくる。拡張現実感というものを表現するためには、やはりアクティブに動き回れる方がよりよいと考えられる。ただし、マーカ型であるARToolKitの特性上、マーカレス型のPTAMに比べユーザーの自由度が低くなってしまうということが挙げられる。

(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頻度が少しの色でもそれを最大として読み込むために、常に何かしらの反応を返してしまう。その結果、挙動が安定しなくなるので感性情報提示装置としては不備があるといえる。

また、色相値が最大の色の検出を行う手法をとっているため感情を強度毎に表現することが出来ないという欠点がある。対策としては、前景となる部分が全体のある一定以下であった場合は無視するといった事が考えられる。

(4) 感情の強度

本研究では感情の強度という概念を無視したそのため、細かな感情表現ができなかった。そこで、感情の強度毎にモーションを作成すればさらに親しみやすいインターフェイスが作成可能であると考ええる。

最終的な目標としてネットワーク・拡張現実感・感性情報処理の融合を実現したい。今回の研究では拡張現実と、感性情報処理の融合として「AR版感性君」を作成した。これにネットワークを組み込むことで例えばチャットシステムが構築可能であろう。

また、「AR版感性君」のシステム自体はプログラムにより作成しているのも機能の拡張を図ることも可能だ。例えば、今回用いた色識別だけでなく、音声認識や、文字認識を用いることでよりバリエーション豊富な感情表現を行える。また、システムを一新しなくてはならないがマーカレス型にすることでよりリアルに近い拡張現実を表現可能であろう。

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]西岡龍太,「主観観測モデル理論を用いたロボットの主観を伴う感情と動作の自律生成」,北九州工業高等専門学校専攻科特別研究論文,2005