

筑波大学大学院博士課程

システム情報工学研究科修士論文

ライフログの閲覧と分析に
人間の表情を活用するシステム

高井 啓

修士（工学）

（コンピュータサイエンス専攻）

指導教員 田中 二郎

2014年3月

概要

「人の顔色を伺うコンピュータ」という概念を提唱し、その一例として、ライフログの閲覧に人間の表情を活用するシステムを提案する。表情は人間の持つ最も基本的なインタラクション手法であり、コンピュータが表情を認識し続けその状況に合わせたインタラクションを行うことができれば、日々の使用に於いて人間の感性に則したシステムができると考えた。そして、その一例として撮りためた表情を元に、ライフログの閲覧を支援するシステムを提案し作成した。提案システムは表情データを利用することで大量のライフログデータの中から表情をキーとしたライフログデータの検索を可能とする。我々はケーススタディとして、PCの画面を記録し続けたライフログデータを、提案システムを使って閲覧し分析を行った。本論文ではその結果を元に、表情を活用したライフログの閲覧の有効性、さらに人の顔色を伺うコンピュータについて論ずる。

目次

第1章 序論	1
第2章 関連研究	2
2.1 ライフログの閲覧	2
2.2 表情を利用したシステム	3
第3章 システム構成	4
3.1 記録部の実装	4
3.1.1 表情データの形式	4
3.2 分析部	5
3.2.1 閲覧ソフトウェア	6
3.2.2 分析ソフトウェア	6
類似度の種類	7
第4章 評価	11
4.1 記録部の評価	11
4.1.1 AU情報のヒストグラムの提示	11
4.1.2 散布図作成プログラム	17
4.1.3 散布図での比較	17
4.2 解析部の評価	26
4.2.1 プログラミングのデバッグを行っている時	29
4.2.2 動画を見て笑いをこらえている時	32
4.2.3 顎に手を当てて考えている時の顔	35
4.2.4 動画を見て大きく笑っている時	38
第5章 考察	41
第6章 結論	42
謝辞	43
参考文献	44

図目次

3.1 システム構成	5
3.2 閲覧ソフトウェア	7
3.3 分析ソフトウェア	10
4.1 ヒストグラム作成プログラム	12
4.2 ヒストグラム AU0	13
4.3 ヒストグラム AU1	14
4.4 ヒストグラム AU2	14
4.5 ヒストグラム AU3	15
4.6 ヒストグラム AU4	15
4.7 ヒストグラム AU5	16
4.8 散布図作成プログラム	17
4.9 X 軸 AU0, Y 軸 AU1	18
4.10 X 軸 AU0, Y 軸 AU2	18
4.11 X 軸 AU0, Y 軸 AU3	19
4.12 X 軸 AU0, Y 軸 AU4	19
4.13 X 軸 AU0, Y 軸 AU5	20
4.14 X 軸 AU1, Y 軸 AU2	20
4.15 X 軸 AU1, Y 軸 AU3	21
4.16 X 軸 AU1, Y 軸 AU4	21
4.17 X 軸 AU1, Y 軸 AU5	22
4.18 X 軸 AU2, Y 軸 AU3	22
4.19 X 軸 AU2, Y 軸 AU4	23
4.20 X 軸 AU2, Y 軸 AU5	23
4.21 X 軸 AU3, Y 軸 AU4	24
4.22 X 軸 AU3, Y 軸 AU5	24
4.23 X 軸 AU4, Y 軸 AU5	25
4.24 プログラミング時のスクリーンショットの例	27
4.25 サーベイ時のスクリーンショットの例	28
4.26 2013/11/18 20:28:29.100 時点の顔	29
4.27 2013/11/18 20:28:29.100 時点のスクリーンショット	30
4.28 2013/11/26 18:28:0.324 時点と似ている表情のシーン分類	30

4.29	2013/11/26 18:28:0.324	時点と似ている表情のシーンの作業と非作業分類 . .	31
4.30	2013/11/26 14:57:39.392	時点の顔画像	32
4.31	2013/11/26 14:57:39.392	時点のスクリーンショット	33
4.32	2013/11/26 14:57:39.392	時点と似ている表情のシーン分類	33
4.33	2013/11/26 14:57:39.392	時点と似ている表情のシーンの作業と非作業分類 . .	34
4.34	2013/11/26 18:28:0.324	時点の顔画像	35
4.35	2013/11/26 18:28:0.324	時点のスクリーンショット	36
4.36	2013/11/26 18:28:0.324	時点と似ている表情のシーン分類	36
4.37	2013/11/26 18:28:0.324	時点と似ている表情のシーンの作業と非作業分類 . .	37
4.38	2013/7/6 22:20:21.898	時点の顔画像	38
4.39	2013/7/6 22:20:21.898	時点のスクリーンショット	39
4.40	22013/7/6 22:20:21.898	時点と似ている表情のシーン分類	39
4.41	22013/7/6 22:20:21.898	時点と似ている表情のシーンの作業と非作業分類 . .	40

第1章 序論

表情は人間が関係を築く為に利用する基本的なインタフェースである。人間は会話や協調作業の際に、表情を用いてコミュニケーションを行う。本研究では、そのような人間の特性を応用したシステムとして、人間の表情を読み取り、その情報をコンピュータのインタラクションに応用するシステムを、「人間の顔色を伺うコンピュータ」として提唱する。

カメラデバイスが安価になってきたこともあり、多くの携帯端末にカメラが搭載されるようになってきている。コンピュータが日常的に表情を読み取り、その情報を役立てるための条件は整いつつある。

しかしながら、人間の表情は個人差が大きい為、万人に等しく適用することができる認識方法を作り出すことは難しいという問題点がある。そこで、あえて認識を行わずに、表情の類似度を用いて過去の大量のデータの中から似ている表情を持ったシーンを取得し、インタラクションを行う仕組みを提唱する。

例えば嬉しい顔をした時に行った操作を学習し、それを繰り返すことができれば、ユーザーにとってよりストレスの無い操作が可能になる。この操作に必要なのは類似度の計算だけであり、「嬉しい」感情を評価する適当な指標さえあれば、学習は可能だと考えられる。

本研究ではそのシステムの実現に向けて、人間の表情を記録し続け、表情の類似度を用いて過去に記録されたデータからシーンを検索してくるシステムを作成し、ライフログの閲覧と分析を行うシステムとして実装する。

ライフログの閲覧と分析を行うシステムにこの手法を応用した理由は、ライフログに関して、大量に記録されたライフログデータの中から効率的にデータを閲覧する研究が行われており、[TYA05, 奥浦 11, 上岡 04, MKK⁺12, FTT13]。ライフログデータの閲覧にまつわる問題を解決するシステムを作成することで、表情の類似度を用いて過去のデータの中から有益なシーンを探して来る手法について検討することができると考えたからである。

第2章 関連研究

関連研究について述べる。ライフログの閲覧に関する問題を解決する研究と、表情を利用する研究について取り上げる。

2.1 ライフログの閲覧

ライフログデータの閲覧や分類を支援する試みが行われてきた。

Datchakorn ら [TYA05] は、ウェアラブルカメラからの動画、マイクロフォンからの音声、モーションセンサー、GPS、皮膚抵抗値、体温、心拍数、等の情報を記録し続けて作成されたライフログに対して、コンテキストに注目してサマリとして抜き出す研究を行った。これにより、会話等、ユーザにとって重要な意味を持つ部分を検索することが可能となった。

牛越ら [牛越 09] は周囲に存在する Bluetooth デバイスからユーザの状態を分類する手法を提案し実験を行った。論文の時点では、手法の提案のみに留まっているが、関連研究としてライフログをあげており、ライフログへ応用することにより、ユーザの行動を分類し、表示することができる。

奥浦ら [奥浦 11] は、周囲に居るユーザが持つ Bluetooth 対応機器の ID を取得し、デバイスタグとして写真に関連付けるシステムを提案し、実装した。このシステムは、写真が取られた時に近くに存在した Bluetooth 対応機器の ID を照らし合わせることで、写真を撮影した時に近くに居たユーザ、また機材の組合せ等の状態の変化を識別することができ、関連性の高い写真をユーザに提示することができる。

上岡ら [上岡 04] は、カメラ画像から物体の色相、反射赤外高輝度、画像重心からの距離から物体を識別し、記録することにより、もの探しを効率化するウェアラブルシステムを作成した。ユーザがこのシステムを利用して、探したい物体を選択すると、最後に物体が記録された時間の映像がユーザに提示される。

Daniel ら [MKK⁺12] は、web カメラ、マイク、震度カメラ、EDA センサ、等のセンサや、File Activity 等の情報を利用したライフログを作成し、その内容を解析することにより、ユーザの感情を解析し提示するシステムを提案し、開発を行った。これにより、ユーザは自分の感情を元にして、ライフログを閲覧することができる。

Doherty ら [DS08] らは、SenceCam による膨大なファイルを、イベント事に自動で区分けする研究を行った。区分けの方法は画素の関係から、大きく場面が変わった部分を検出することで行っている。

福本ら [FTT13] は、フォトインタラプタを利用して目元の筋肉の動きを取得するメガネ状の

デバイスを作成し、その情報を元に、ウェアラブルカメラによるライフログビデオから、ユーザが笑顔の時に見えていた部分を取得し、提示する方法を提案した。この手法では目元の筋肉を判別に用いているので、数種類の表情の変化しか取得できないが、本研究は Kinect を用いており、人間の浮かべる全ての表情について扱うことが可能である。

2.2 表情を利用したシステム

表情を利用してシステムを作成する研究について述べる。

Javier ら [HHDP12] は野外にカメラを設置し、そのカメラに写る人間の表情を認識し、屋外における表情のマップを作成した。

平山ら [平山 07] は顔の特徴点を取得し、さらにタイミング構造と合わせて解析するというシステムを考案し、同じ笑いでも「作り笑い」と「自然な笑い」の識別が可能になったと述べている。一方、本研究は表情を識別した後、それをライフログと関連付けて利用する。

中村ら [中村 05] は、e-learning システムを用いた非同期型の遠隔教育において、学習者の顔を記録し、重畳表示するシステムを作成した。このシステムにより、ユーザは学習者の理解度を把握することができる。この研究では表情の画像を記録し表示するものであり、表情の解析は特に行っていない。

Joho ら [JJVS09] や古川ら [古川 13] の研究では、カメラで動画閲覧中の人間の表情を取得し、それを利用して動画のダイジェストを作成する研究を行っている。

岩松ら [岩松 13] は、携帯端末のインカメラで撮影した画像から表情を認識し、ソーシャルグラフの作成に活用する提案を行った。

紙谷ら [紙谷 97] は、ユーザの状態を表情から推測し、適切なレスポンスを返す為に、困り顔を判別する方法について動作解析を行った。本研究では、レスポンスは特に返さないが、困り顔だけでなくユーザの表情全てを認識し、ライフログの閲覧に役立てることが出来る。

いずれの研究も表情を識別するために様々な手法を利用しているが、類似度を用いてシーンを分類するアプローチは存在せず、本研究とは異なっている。

第3章 システム構成

システムを構成するにあたり，データを記録する記録部と，そのデータを利用して情報の提示を行う分析部に分けて設計を行った．図 3.1 にシステム構成を示す．

記録部はシステムに必要な情報を記録する部分であり，ライフログデータや表情ログを記録し続ける．分析部は撮りためたデータを閲覧し，どのようなタイミングでどのような表情の変化があったのかを確認する閲覧ソフトウェアと，記録した表情を解析し，表情の類似度から過去のライフログデータを検索し，似ていると考えられるシーンをランキング形式で提示する分析ソフトウェアから成り立つ．

ユーザはまず記録部を利用してデータを撮りためる．次に閲覧ソフトウェアを利用して，記録されたデータの中から自分が興味をもったシーン（以下基準点）を探し出す．その後分析ソフトウェアを利用して，その基準点の表情に似ている表情を持ったシーンを記録されたデータの中から探し出し，ランキングを作成する．

3.1 記録部の実装

記録する情報は顔画像と表情データ，ライフログデータの3つである．顔画像と表情データは Kinect とそれに付随するライブラリを利用して記録する．将来的に携帯端末に搭載されたカメラやラップトップ PC に搭載されたカメラでも動作するようになれば良いのだが，現時点では形態端末やラップトップ PC の計算速度が不十分であり，別途デバイスを用いることとした．さらにライフログデータとして，PC の画面のスクリーンショット（以降，画面ログ）も取得する．画面ログを採用した理由はユーザが何を見ていたかが直ぐに解る為，実験やデバッグに適していると考えたからである．

記録部は C # で記述されている．また記録部は表情データを概ね 1 秒間に 10 回のペースで，顔画像と画面ログを概ね 1 秒間に 1 回のペースで記録する．

3.1.1 表情データの形式

表情データは KinectSDK で提供される AU (Animation-Unit) をそのまま利用している．AU は KinectSDK で表情を表す為に使われる値であり，AU0 から AU5 の 6 種類が存在し，それぞれ表情の部位に対応している．表 3.1.1 に AU と表情の対応を記す．AU 情報の値は -1 から 1 までの範囲を取り，何も表情を浮かべていない状態（以下無表情）の時には 0 の値と取るように設定されている．

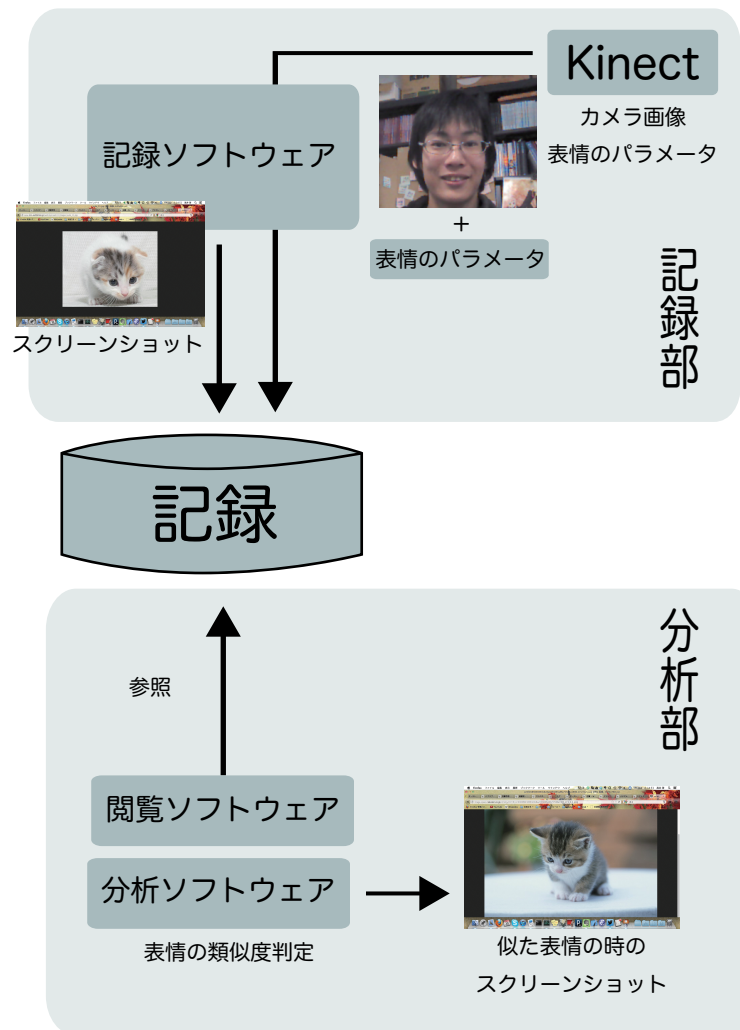


図 3.1: システム構成

これらの情報を 6 次元のベクトルとして考えた時のノルム（大きさ）と，取得時の UnixTime を付け加えて，CSV 形式で保存する．ノルムを付け加えた理由は，3.2.2 章で解説するコサイン角度の類似度の計算にこの値を利用することで計算量を減らす為である．

3.2 分析部

取りためたライフログデータの閲覧や，表情を元に似ている表情のシーンを選び出して分析を行う部分である．

AU0	Upper Lip Raiser	上唇の上がり方を表す. 最も上がってる状態が 1, 最も下がってる状態が-1.
AU1	Jaw Lowerer	顎の下がり方を表す. 1 で口を大きく開けた状態. -1 が閉まった状態 (但しこれは殆ど 0 の状態に近い).
AU2	Lip Stretcher	唇の広がり方. 1 で十分に広がっている状態 (joker's smile と呼称されている) -0.5 でふくれっ面のように丸い状態. -1 でキスをするときのように尖った状態.
AU3	Brow Lowerer	眉毛の下がり方. 1 で限界まで下がった状態. -1 で完全に上がった状態.
AU4	Lip Corner depressor	口角の下がり方. 1 でとても悲しい, しかめっ面. -1 で笑顔.
AU5	Outer Brow Raiser	上眉の外側の上がり方. 1 で驚いた表情. -1 で悲しい表情.

表 3.1: AU 情報の詳細

3.2.1 閲覧ソフトウェア

閲覧ソフトウェアは, 今まで撮りためたライフログデータを閲覧することができるソフトウェアである. ユーザはこのソフトウェアを利用して, ライフログデータの中から自分が探し出したい特徴的な点 (以下, 基準点) を選び出す.

閲覧ソフトウェアのスクリーンショットを, 図 3.2 に示す. 表情データがグラフで表示され (図 3.2 上) またその下に, スライダーで指定された時点での表情 (検索キー) (図 3.2 左下) と, その表情に似た表情 (以降, 類似画像) が表示される (図 3.2 中央下). 類似画像はデバッグのために付け加えられている. ここで利用する類似度は 3.2.2 で解説する類似度のどれにも変更が可能であるが, 図 3.2 では標準ユークリッド距離を用いている. また, 別画面でその時のスクリーンショットも閲覧することが可能である. ユーザはこのソフトウェアを利用して, ライフログの中から, 基準点を選び出す.

3.2.2 分析ソフトウェア

分析ソフトウェアは記録部によって保存されたデータを元に, 似ている表情のシーンを検索しランキングを生成することができるソフトウェアである. 分析ソフトウェアの外観を図 3.3 に示す. 図 3.3 の画面上部には基準となる顔画像が, 図 3.3 の画面下部にはそれと類似している顔画像が表示されている. それぞれの顔画像の右側には対応するタイミングの画面ログが表示されている. 画面上部のスライダーを動かすことによってランキング順に顔画像を切り替える事ができる. なおランキングに関して今回は 100 位までのランキングを生成して用いているが, この値は自由に変更することができる.

取得するシーンであるが, 時間軸で近い位置の値は除いて取得する. これは, 近い位置のデータは似ている為に, 類似していると判断される画像がほぼ全て基準点の周りに集中する



図 3.2: 閲覧ソフトウェア

からである。基準点から時間軸上でどれくらい離れた位置から取得を行うかは変更が可能だが、今回は表情データが 1000 フレーム以上（fps により 100 秒以上）離れた部分から取得を行った。

また、分析ソフトウェアはその結果を HTML ファイルとして出力する機能がついている。この HTML ファイルには顔情報とスクリーンショット、類似度等がランキング順に表示されている。

類似度の種類

分析ソフトウェアで利用する類似度は以下の 3 つである。それぞれ、各 AU 情報を要素として持つ、6 次元のベクトル $\vec{X}$ と $\vec{Y}$ とみなして計算を行っている。また、以下、ベクトル $\vec{X}$

と $\vec{Y}$ の各要素を

$$\vec{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_6 \end{pmatrix}, \vec{Y} = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_6 \end{pmatrix}$$

とおく．

二点間で得られた表情のコサイン角度 $\vec{X}, \vec{Y}$ の成す角度を θ とおいて，その角度を類似度を用いる． $\cos \theta$ は式 (3.1) により計算することができる．

$$\cos \theta = \frac{\vec{X} \cdot \vec{Y}}{|\vec{X}| |\vec{Y}|} \quad (3.1)$$

$$\cos \theta = \frac{\sum_{i=1}^6 X_i Y_i}{\sqrt{\sum_{i=1}^6 X_i^2 \sum_{i=1}^6 Y_i^2}} \quad (3.2)$$

類似度は-1 から 1 の間を取り，1 に近づくほど類似性が上がり，最も類似していない場合は-1 の値を取る．初期はこの類似度を利用して計算を行っていたが，ベクトルの成す角度が似ているだけで似ていると判断されてしまうので，無表情の顔の判別が非常に難しいという欠点があった．ノルムで場合分けをする，という工夫により解決出来る可能性もあるが，ユークリッド距離に切り替えてからはそちらを利用している．

ユークリッド距離 $\vec{X}$ と $\vec{Y}$ の間のユークリッド距離は，

$$\sqrt{\sum_{i=1}^6 (X_i - Y_i)^2} \quad (3.3)$$

と表すことが出来る．この類似度は 0 に近づくほど類似性が上がり，もっとも似ていない時は $2\sqrt{6}$ の値を取る．

標準ユークリッド距離 $\vec{X}$ と $\vec{Y}$ の間の標準ユークリッド距離は，

$$\sqrt{\sum_{i=1}^6 \left(\frac{X_i - Y_i}{\sigma(i)} \right)^2} \quad (3.4)$$

と表すことが出来る．ただし， $\sigma(i)$ = ベクトルの要素 i 全体の標準偏差である．

ユークリッド距離と殆ど変わらないが，各ベクトルの要素の差を標準偏差で割る部分が異なっている．ユークリッド距離は，大きく変化する要素と殆ど変化しない要素があった場合，大きく変化する要素に引きずられて値が大きく変わってしまうが，この操作によりその影響を減らす事ができる．

この類似度はユークリッド距離と同じ様に 0 に近づくほど類似性が上がり，もっとも似ていない時は $2\sqrt{6}$ の値を取る．4.1 章によると，今回取得できたデータはデータによって変化しやすいデータと変化しにくいデータが存在しているということが解る．よって今回の実験ではこの標準ユークリッド距離を用いた．

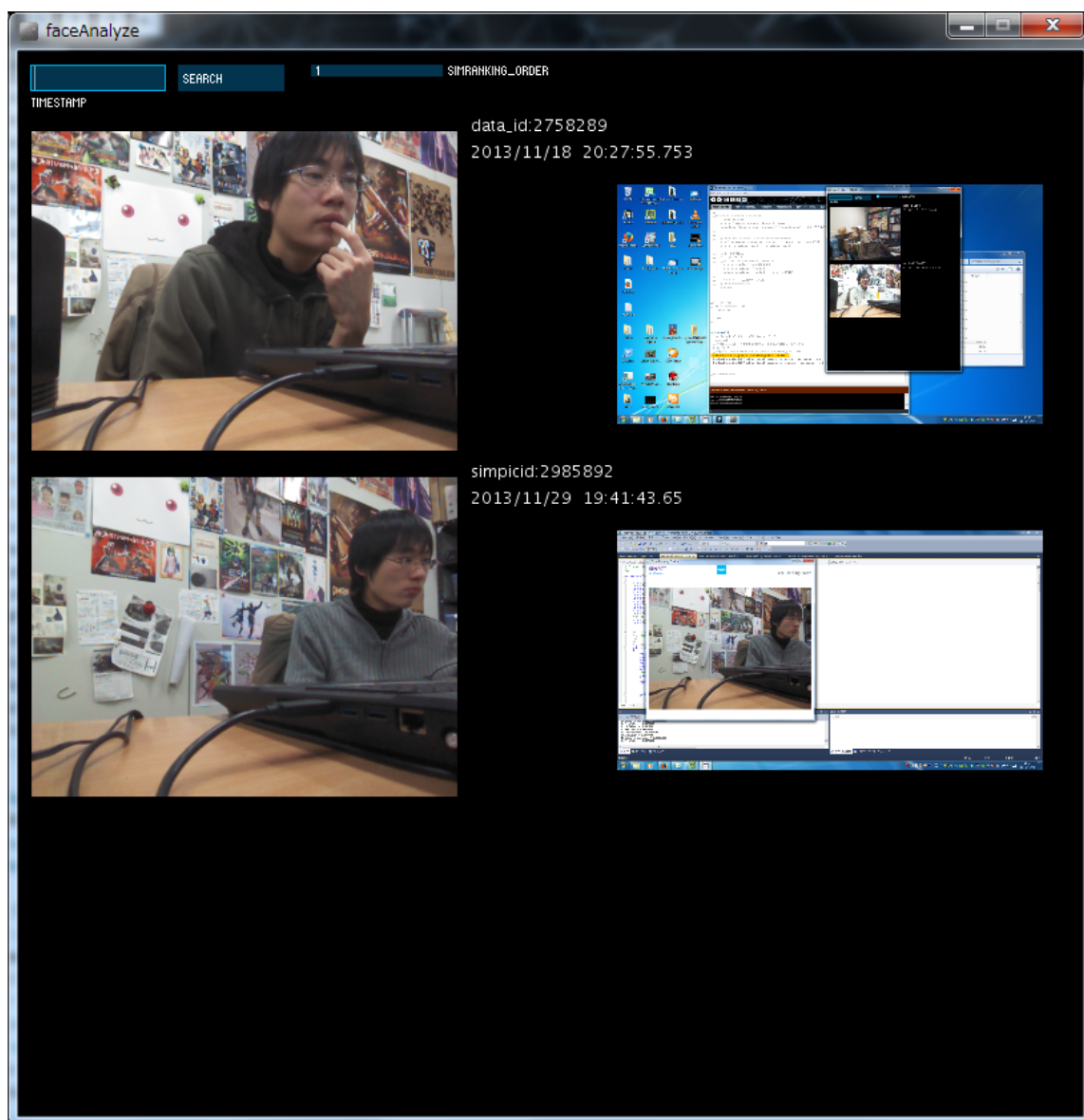


図 3.3: 分析ソフトウェア

第4章 評価

撮りためたデータを元に評価を行った。この評価の目的は、表情の類似度を用いて過去の情報から似ているシーンを取得する手法の有効性について検証することにある。

4.1 記録部の評価

記録部を利用して取得したデータに関して評価を行う。利用したデータは2013年1月07日10時58分から2013年12月23日15時59分までであり、その内容は、筆者が解析部のプログラミングをしている時や、PCを使用して動画又はウェブサイトを閲覧している時、また論文のサーベイを行っている時のデータである。閲覧した動画には、映画のDVDの他、ニコニコ動画で配信されていたコメント付き動画も含む。なお、実装の都合上、2013年6/25日19時59分以前にはスクリーンショットが取得できていない。スクリーンショットは無いものの、表情データのデバッグの為に、映画を何本か見てその表情を取得している。

この撮りためたデータの傾向を探るために、「ヒストグラム作成プログラム」と「散布図作成プログラム」の2種類のプログラムを作成した。これはデータの分布に極端な偏りがあった場合、提案手法の実現に困難が生じると考えられた為である。以下に説明とそのプログラムを使って描画したグラフについての考察を記す。

4.1.1 AU情報のヒストグラムの提示

各AU情報のヒストグラム作成プログラムの全体図を図4.1に示す。

ヒストグラム作成プログラムは、指定したビンの幅（横軸の要素数）¹でヒストグラムを作成して表示することができる。今回はビンの幅を1000に設定して描画を行った。これは分布の形が予想出来なかった為、取り敢えずの値を設定したものである。表示するヒストグラムの縦軸のスケールは変更が可能である。今回は全てのデータが画面内に位置するようにスケールを調節し、同スケールでヒストグラムの描画を行った。描画するヒストグラムはチェックボックスを利用することで切り替えることができる。図4.1では全てのヒストグラムを一度に表示しているが、実際の考察ではAU情報1つ1つに個別のヒストグラムを表示した物を用いる。

図4.2から4.7にヒストグラムを示す。今回グラフ描画に使ったデータの総数はデータの合計数は3210648である、また各AUデータの標準偏差を表4.1.1に示す。

¹ヒストグラムはデータをある一定の数で区切って積み上げたものだが、この区切る個数をビンと言う

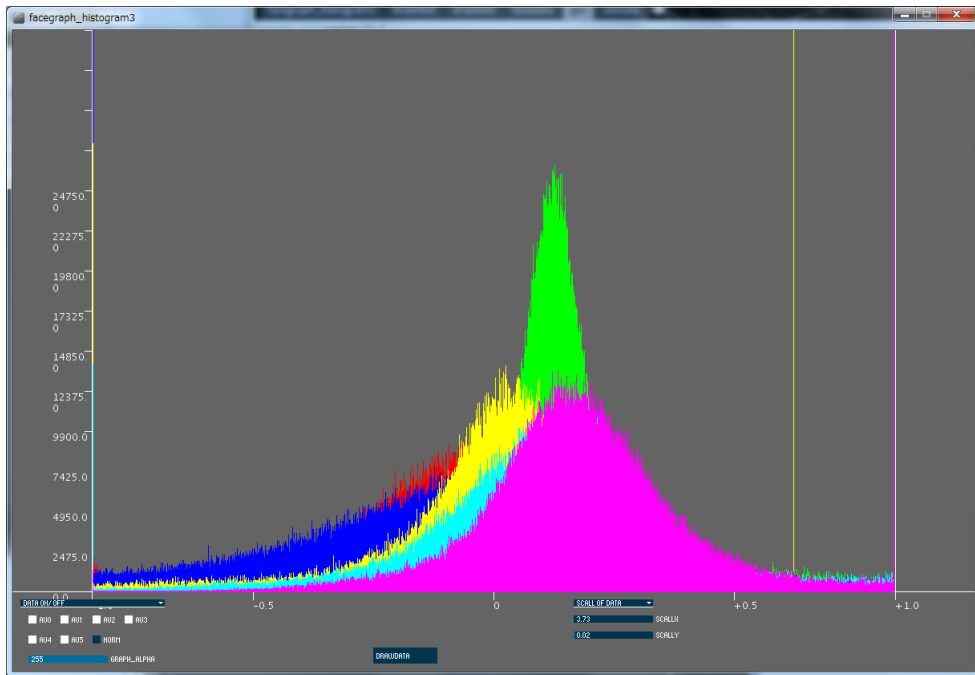


図 4.1: ヒストグラム作成プログラム

AU0	0.4177593169003923
AU1	0.22384124299919123
AU2	0.4331692044549313
AU3	0.30689714190159423
AU4	0.3522590788949453
AU5	0.2979823079888661

表 4.1: 各 AU 情報の標準偏差

全体的にどの分布も山型の分布を取る。すなわち、データが最もよく現れる点は1箇所に限定され、筆者の顔は概ねその部分を中心として分布していることが解る。しかしながら、分布の中心は正確に0に来ているわけではなく、例えばAU0（図4.2）は全体に負の方向に偏っている。また、AU1（図4.3）、AU3（図4.5）、AU4（図4.6）AU5（図4.7）は明らかに0から正の方向に少しずれた部分に分布の中心がある。これは筆者の表情が標準状態で予め定義された無表情から少しずれた場所に位置している為と考える事ができる。

標準偏差を見ると、標準偏差がもっとも少ないAU1（図4.3）は、狭い位置にデータが集中しており、標準偏差通りの分布となっている。AU1は顎の下がり方を表しており、他の表情に比べて口元は変化しにくいことが解る。

AU3（図4.5）は、今回の実験環境ではある一定の値より大きな値を取ることがない。これは眉毛の下がり方であり、ある一定の値より下には筆者の眉毛が下がらない事に起因すると

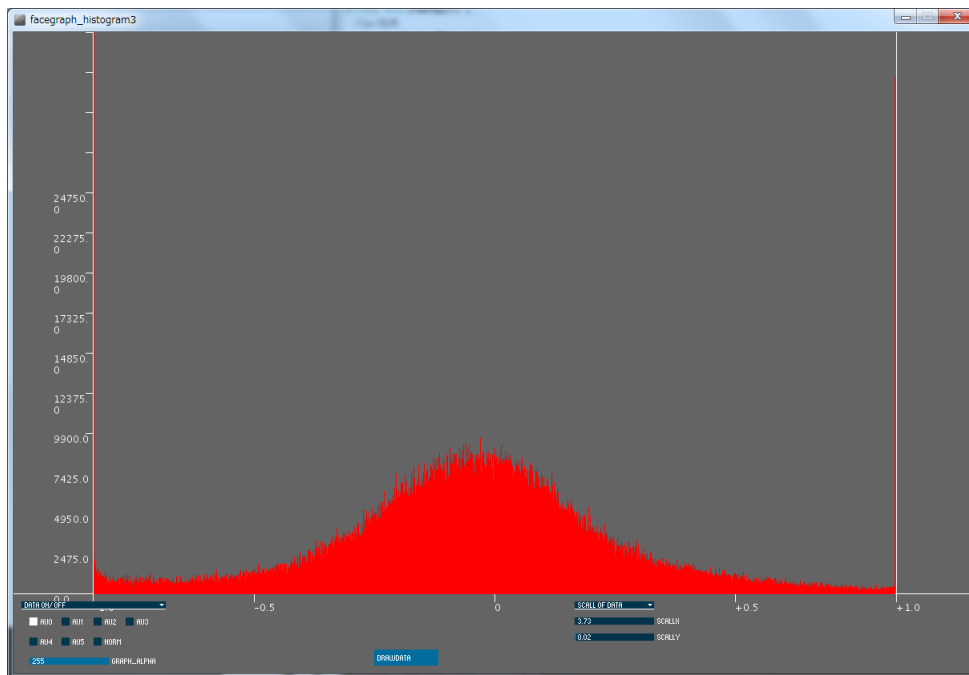


図 4.2: ヒストグラム AU0

考えられる.

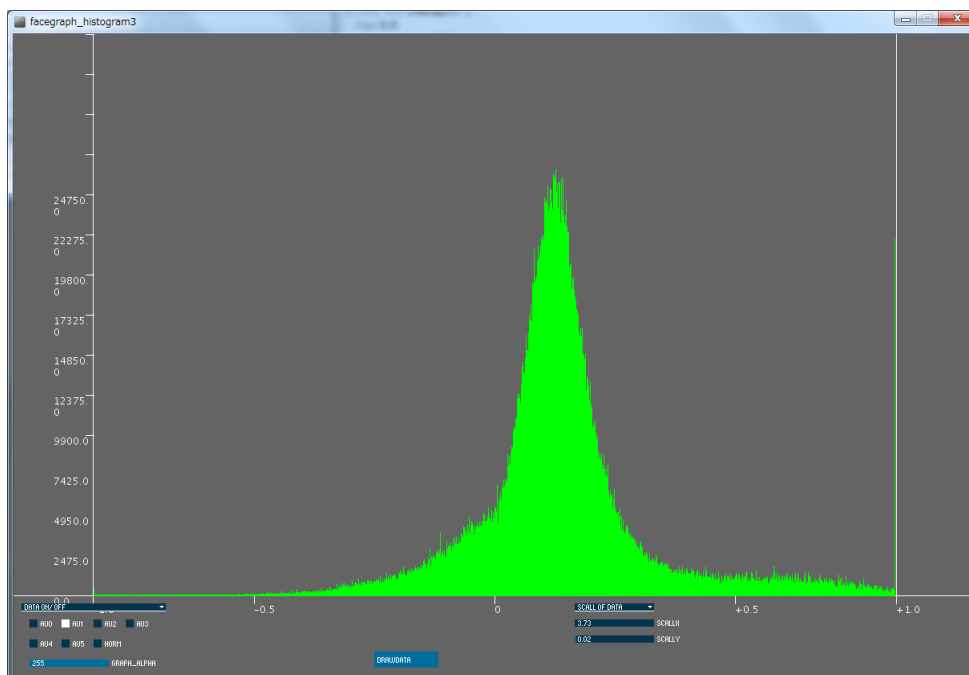


図 4.3: ヒストグラム AU1

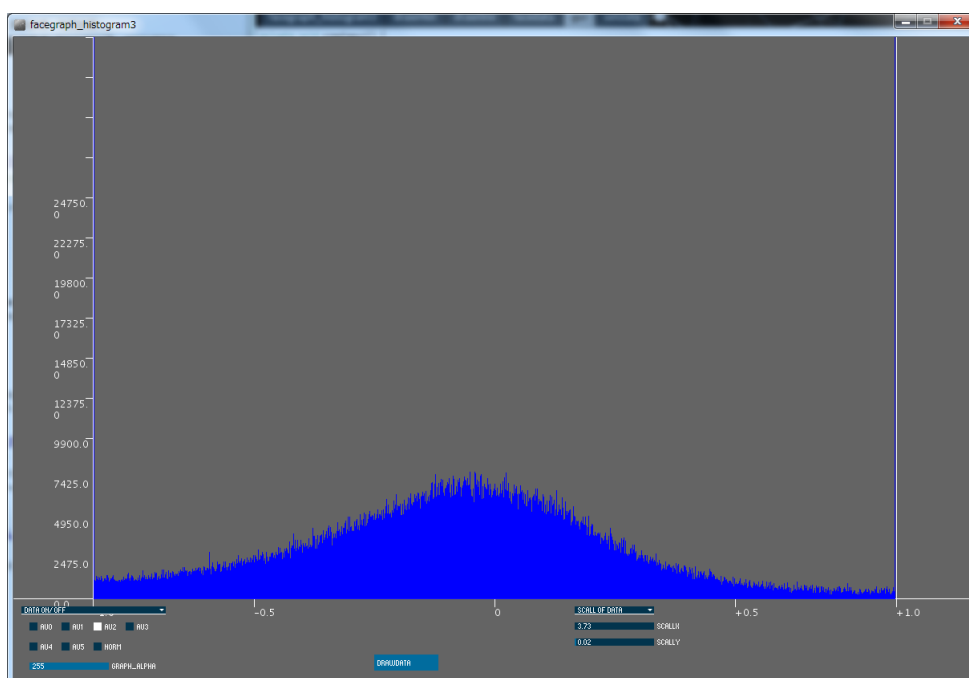


図 4.4: ヒストグラム AU2

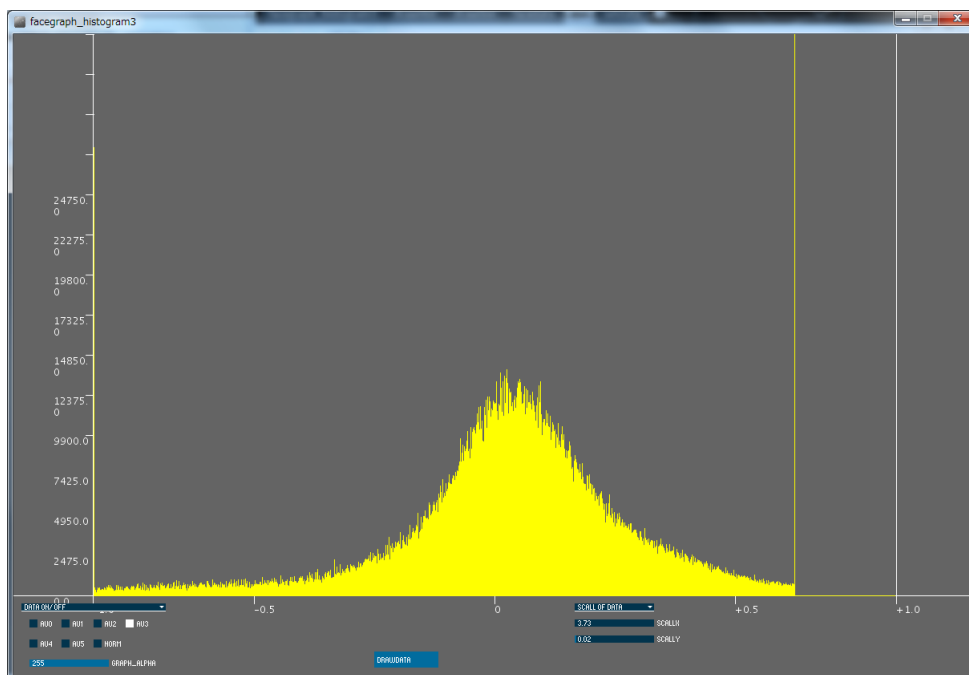


図 4.5: ヒストグラム AU3

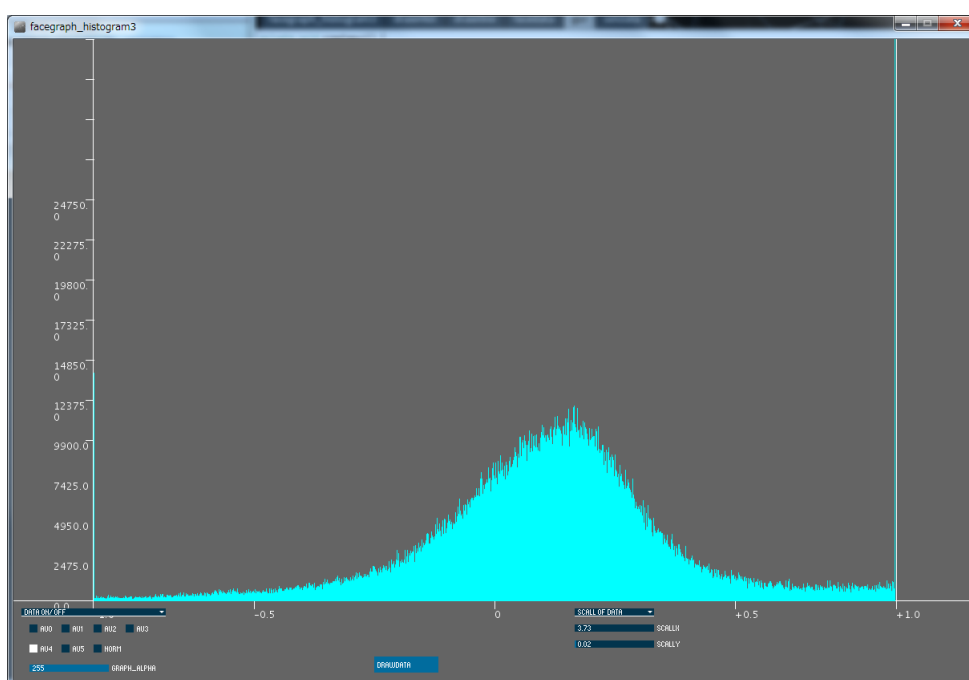


図 4.6: ヒストグラム AU4

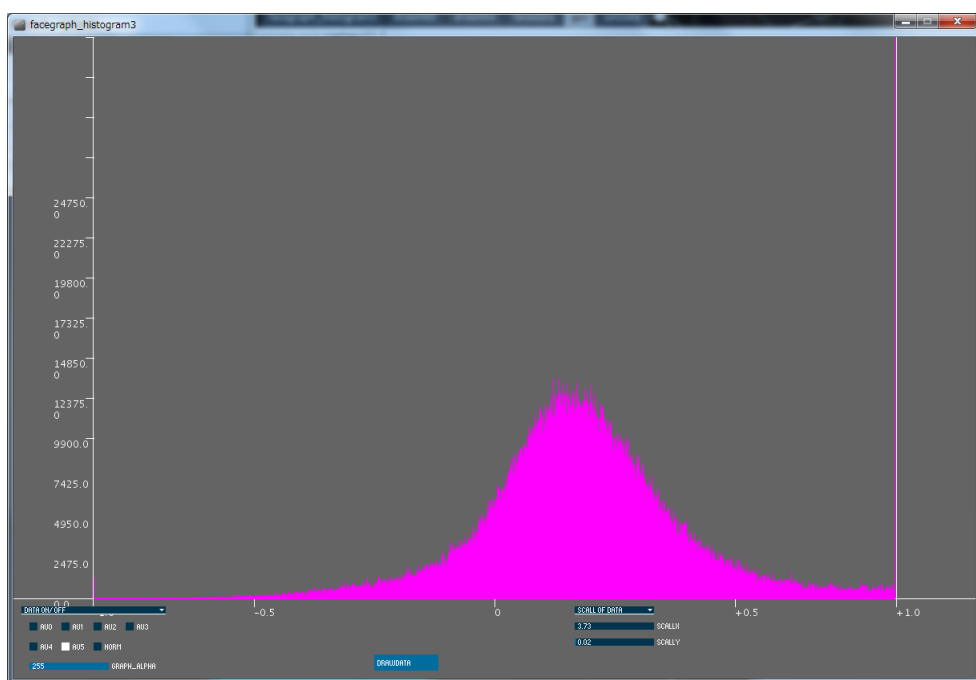


図 4.7: ヒストグラム AU5

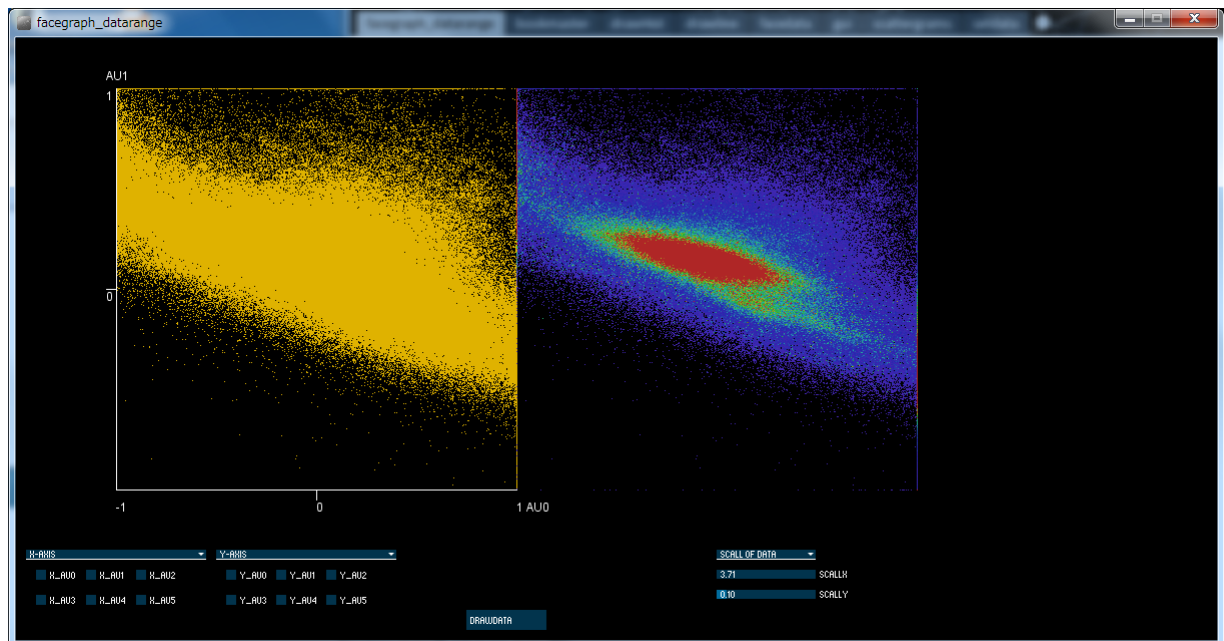


図 4.8: 散布図作成プログラム

4.1.2 散布図作成プログラム

散布図作成プログラムの外観を図 4.8 に示す。

散布図作成プログラムは X 軸と Y 軸に対応する AU 情報を指定すると、散布図を作成するプログラムである。図 4.8 には 2 つの散布図が表示されている。左側の散布図はその位置にデータが一つでも存在すればドットを描画する様にして生成したものである。対して右側のグラフはその位置に存在するデータの数によって色分けを行った散布図である。

色分けは、まず 0 から 255 の範囲で色相を指定し、 $(180 - \text{その位置に存在するデータの数})$ を色相のデータとして利用している。この計算式を適用すると、データ数が最も少ない状態が青として表示され、データ数が多いと赤として表示される。

4.1.3 散布図での比較

散布図作成プログラムを利用して作成した散布図を図 4.9 から図 4.23 に示す。

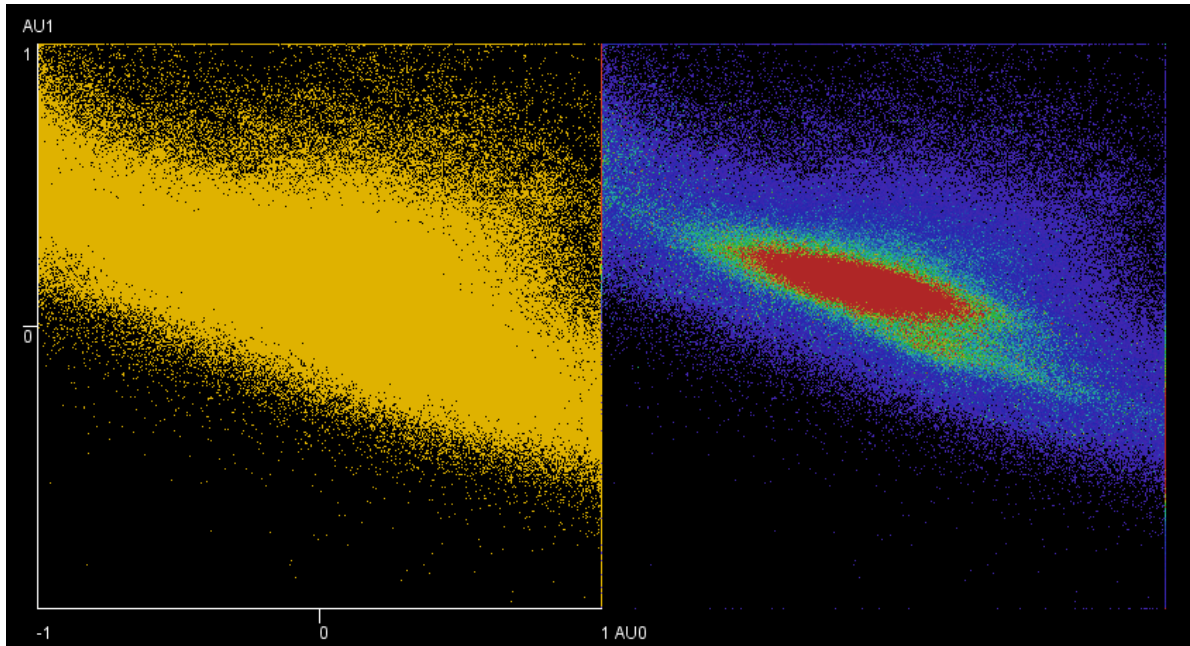


图 4.9: X 轴 AU0, Y 轴 AU1

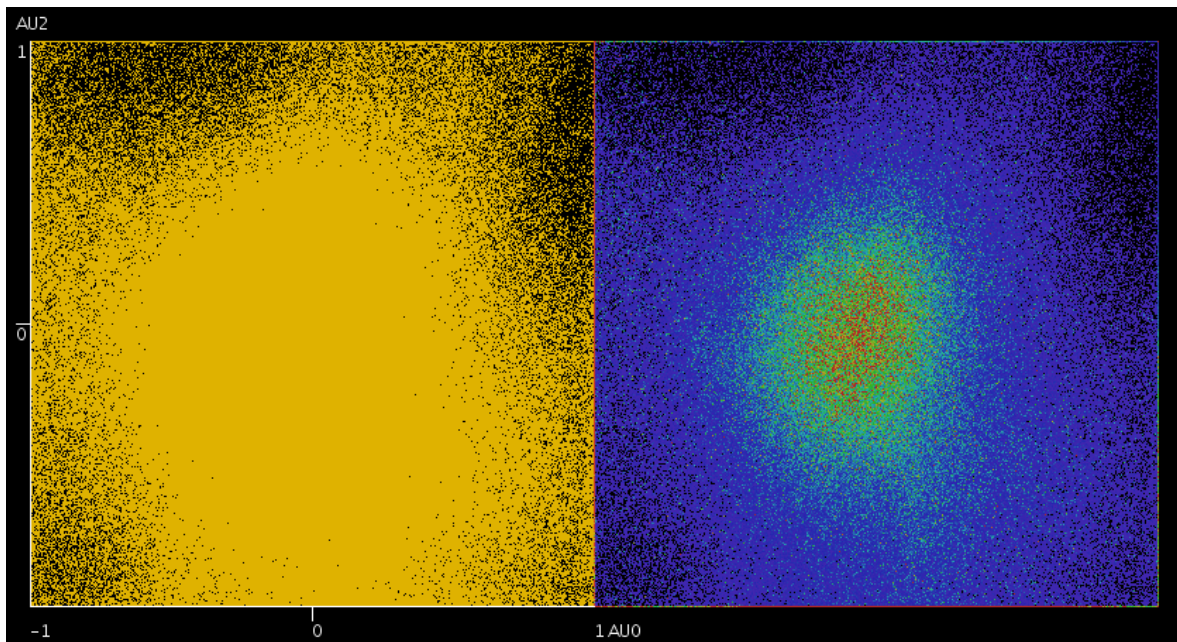


图 4.10: X 轴 AU0, Y 轴 AU2

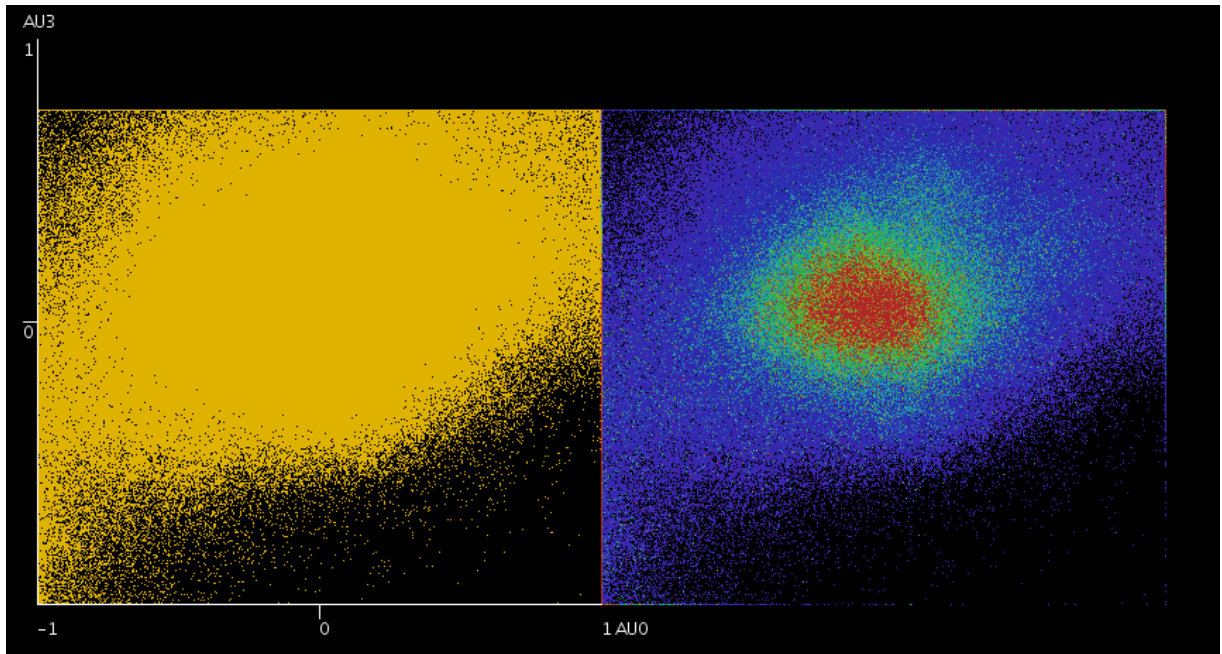


图 4.11: X 轴 AU0, Y 轴 AU3

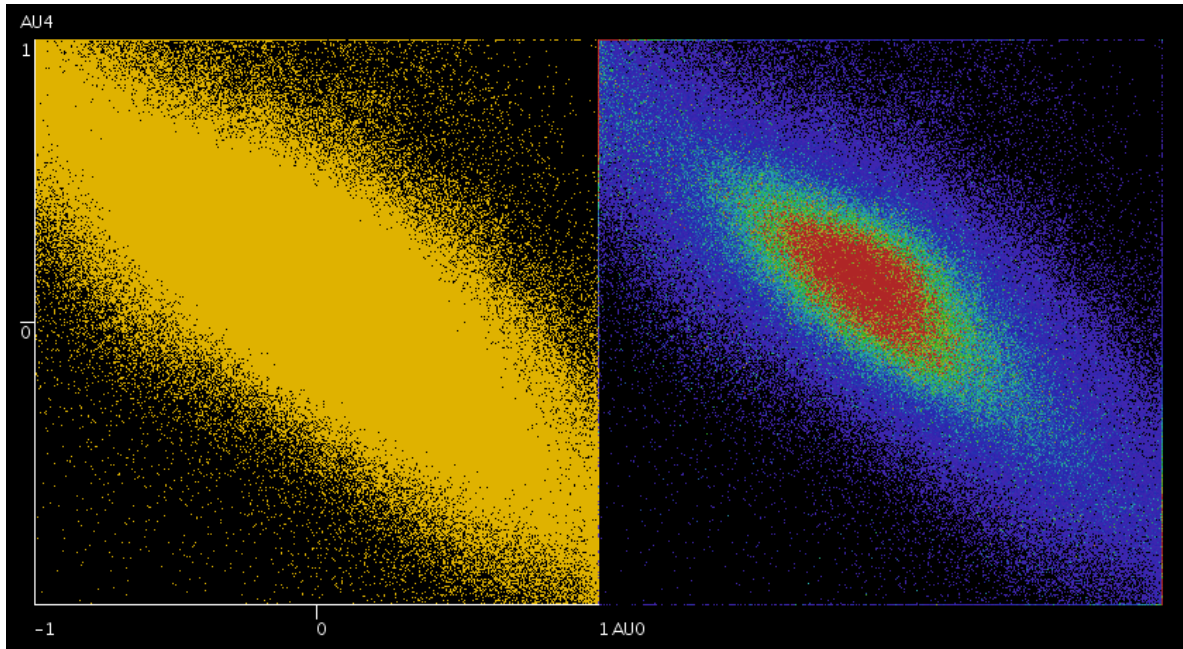


图 4.12: X 轴 AU0, Y 轴 AU4

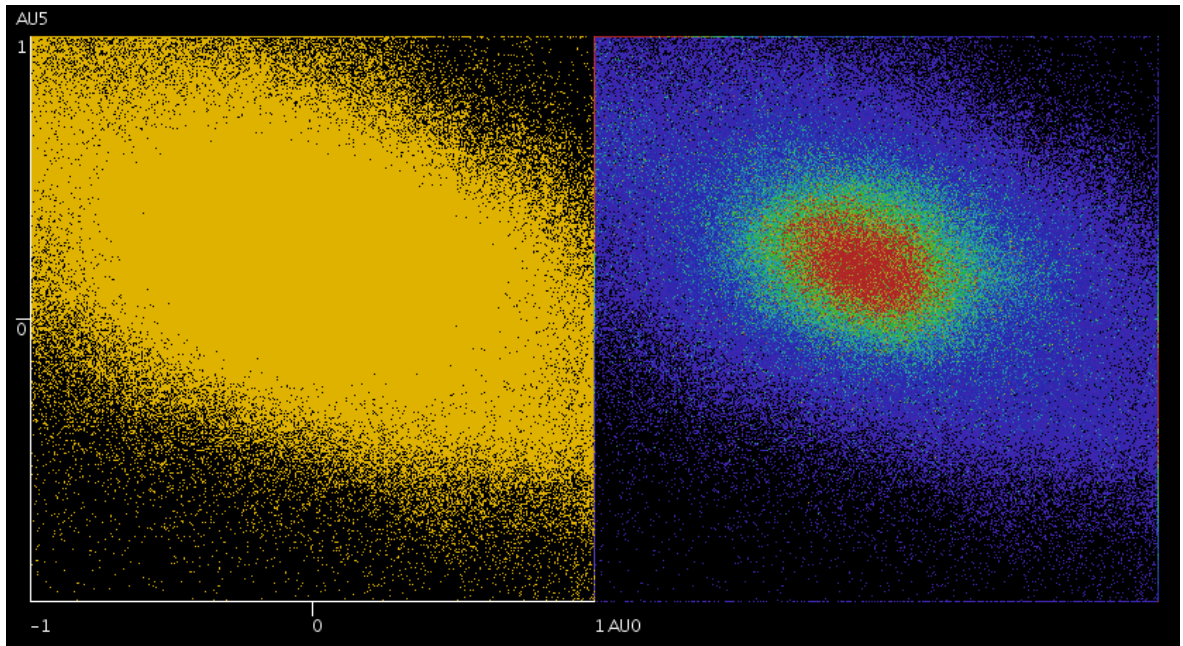


图 4.13: X 轴 AU0, Y 轴 AU5

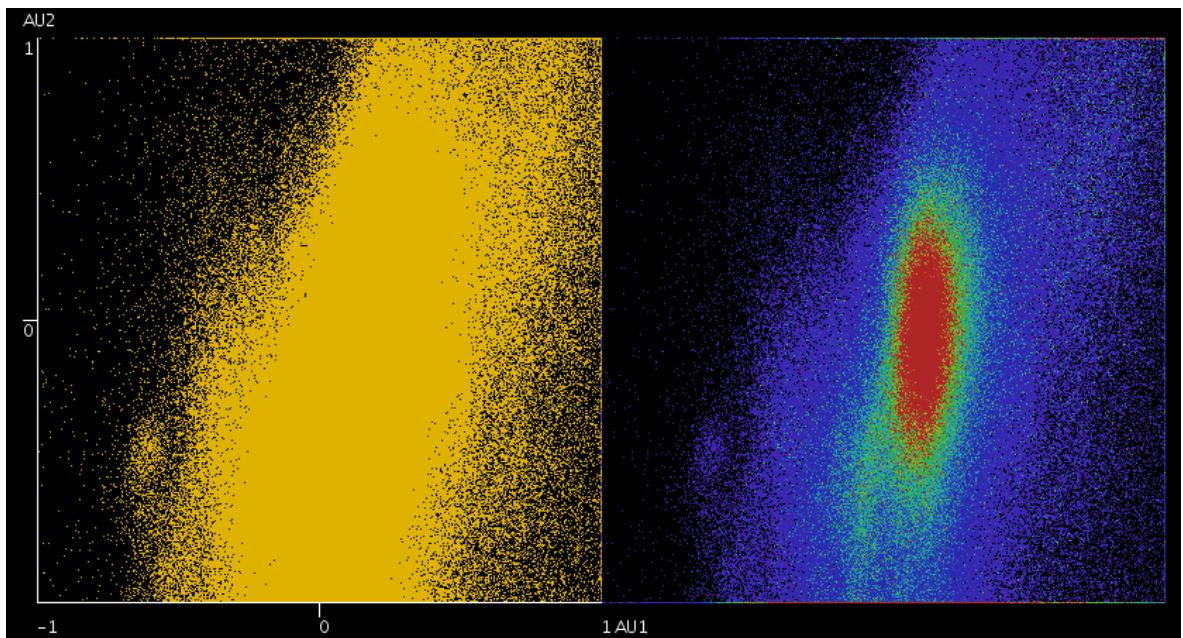


图 4.14: X 轴 AU1, Y 轴 AU2

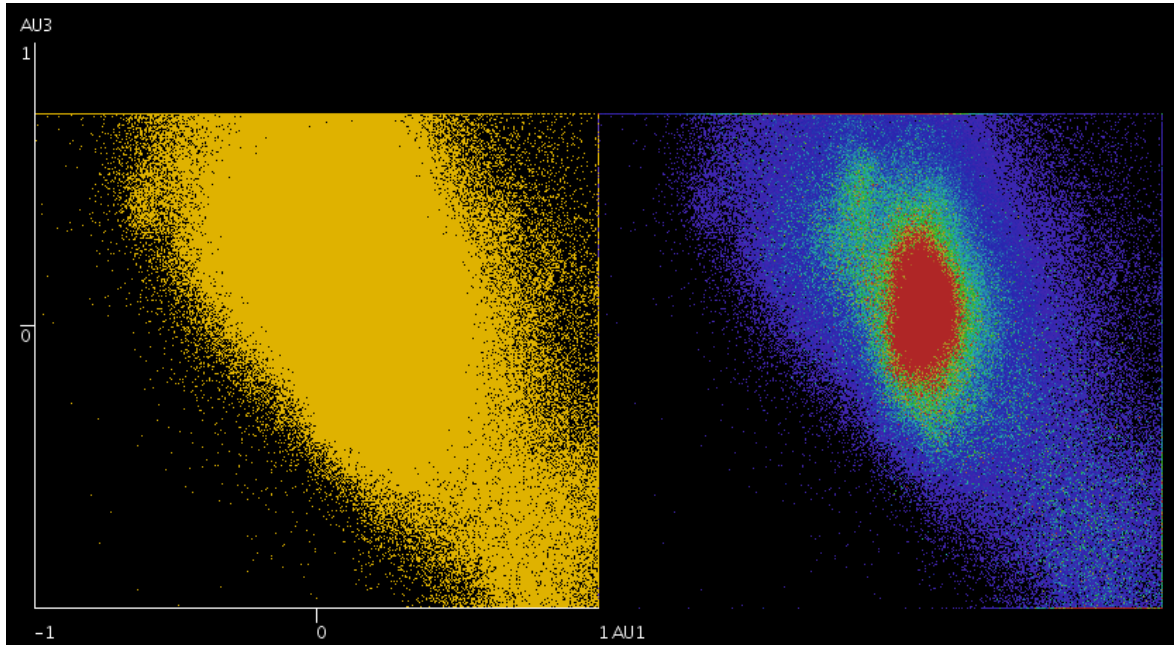


图 4.15: X 轴 AU1, Y 轴 AU3

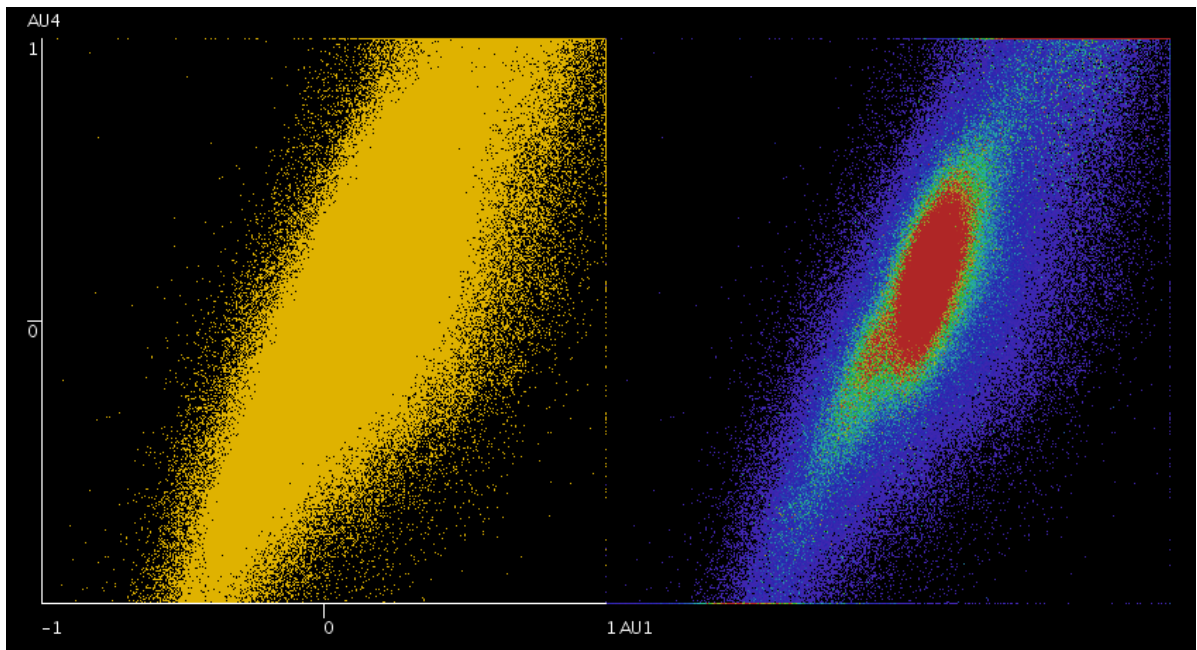


图 4.16: X 轴 AU1, Y 轴 AU4

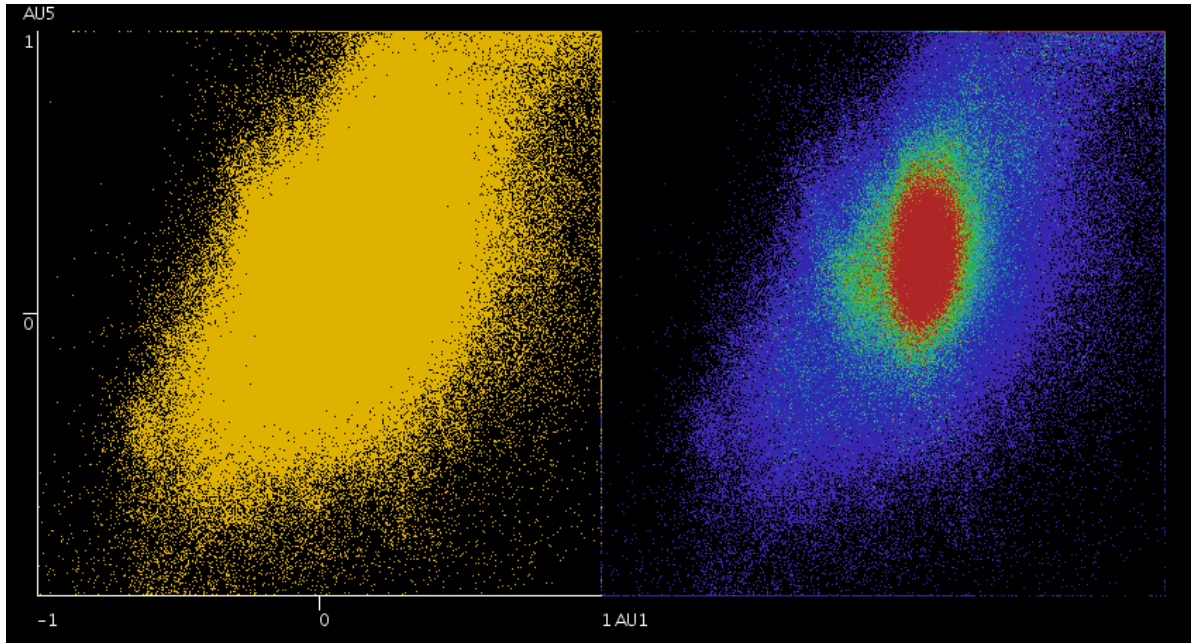


图 4.17: X 轴 AU1, Y 轴 AU5

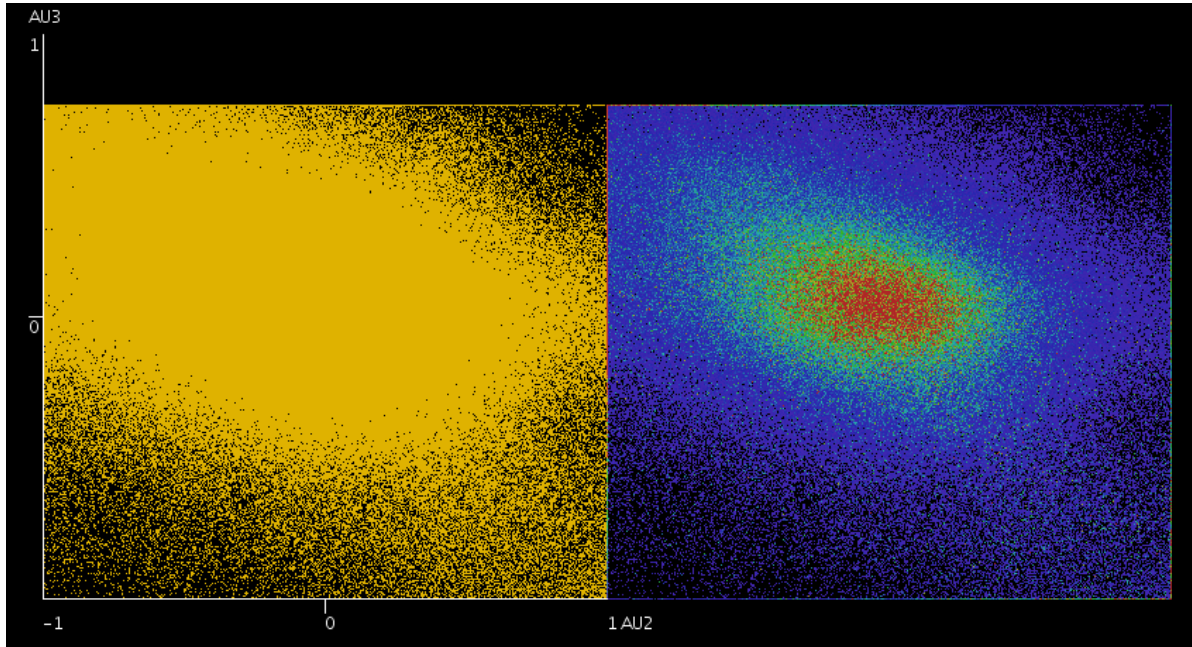


图 4.18: X 轴 AU2, Y 轴 AU3

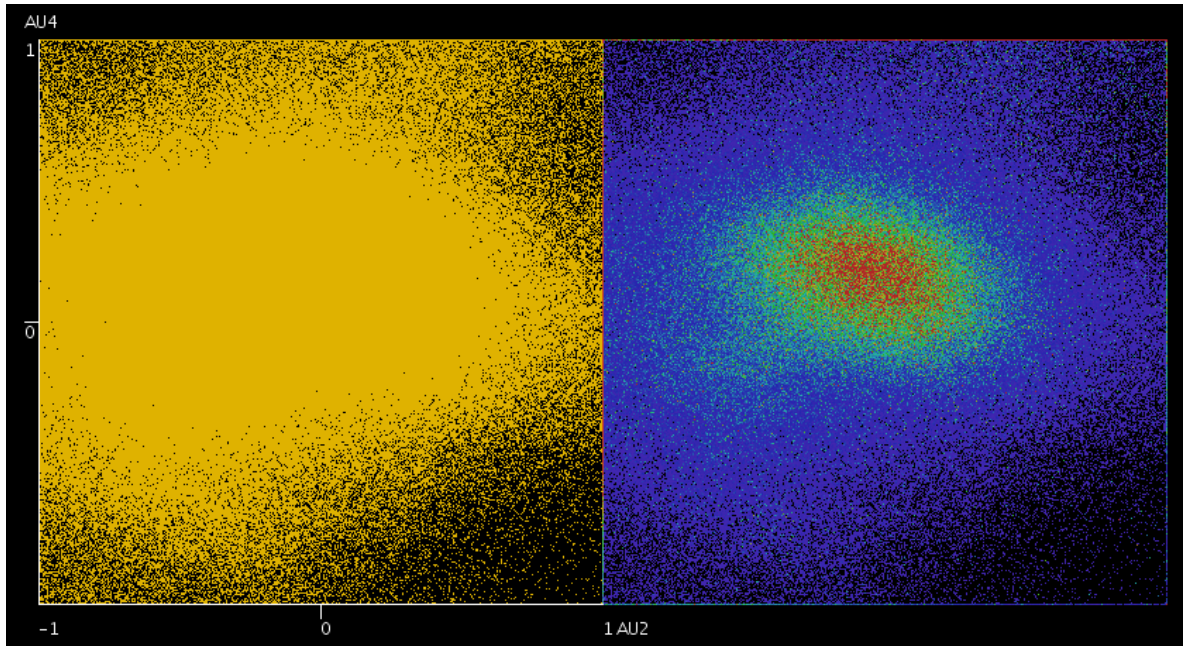


图 4.19: X 轴 AU2, Y 轴 AU4

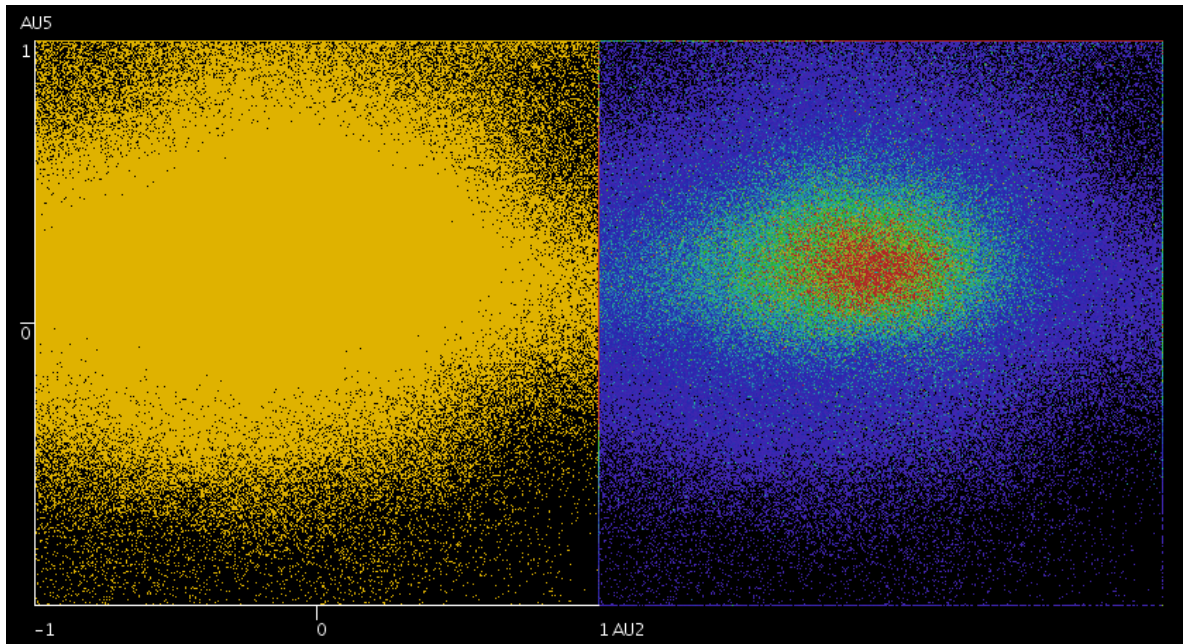


图 4.20: X 轴 AU2, Y 轴 AU5

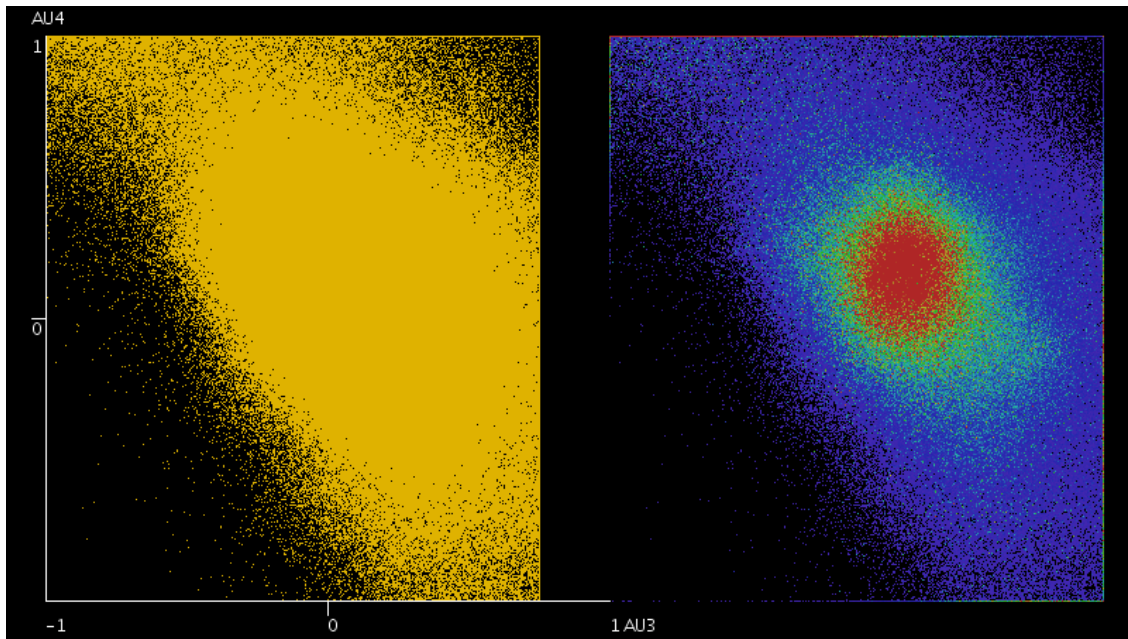


图 4.21: X 轴 AU3, Y 轴 AU4

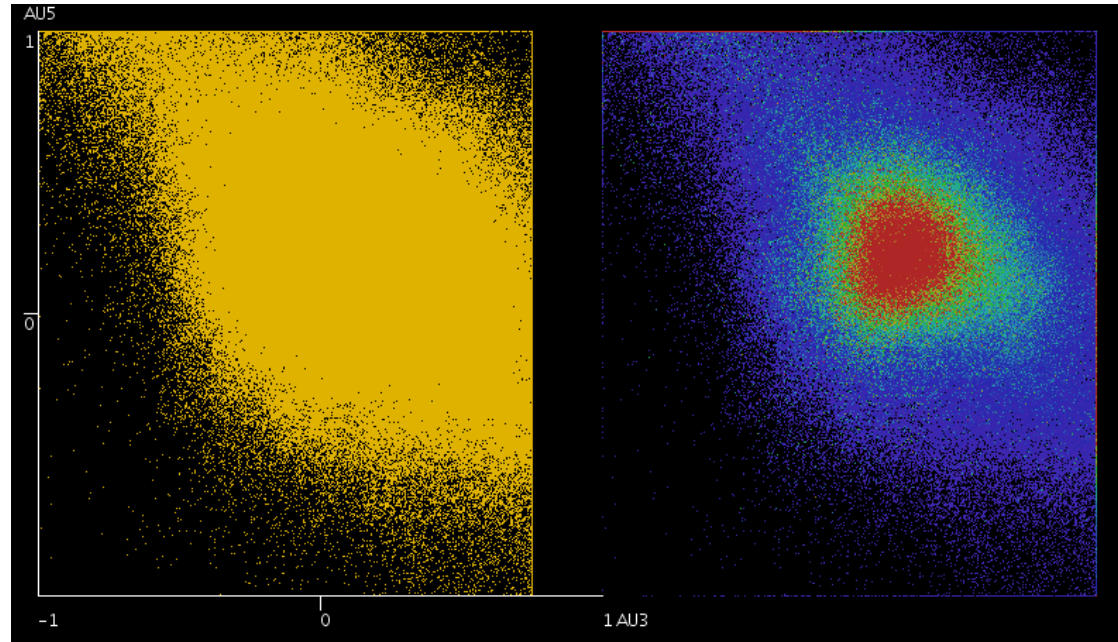


图 4.22: X 轴 AU3, Y 轴 AU5

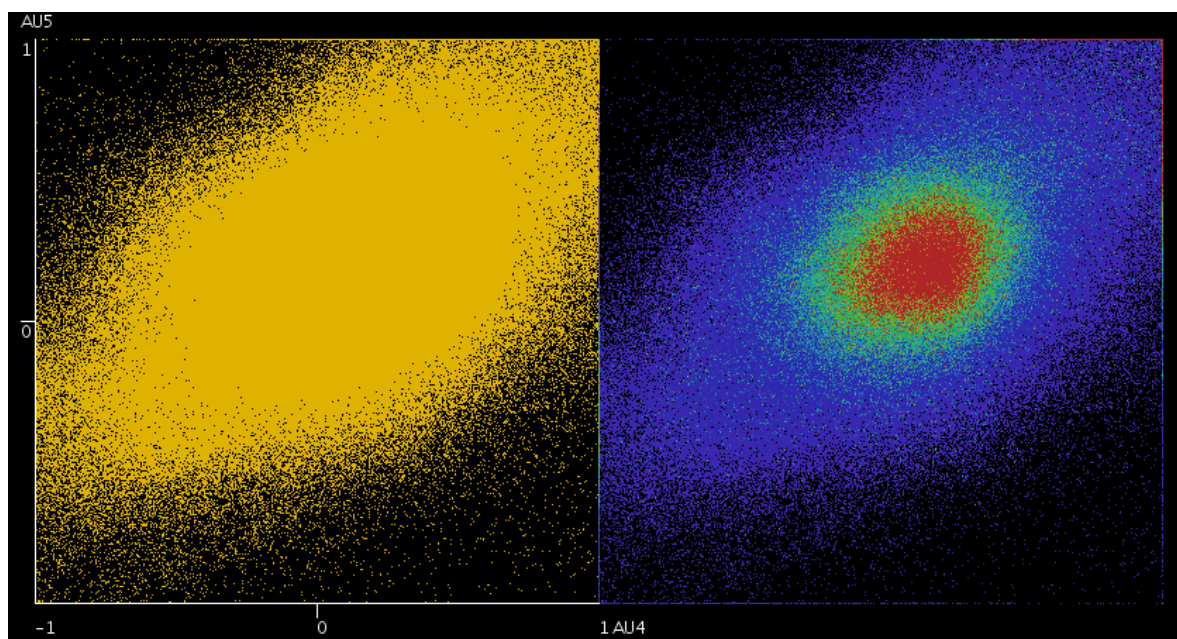


图 4.23: X 轴 AU4, Y 轴 AU5

前提として、AU 情報は無表情の時に 0 に近づくように設計されている。またヒストグラムの結果より、大体 0 付近に集中する場所があるということが解るのでその地点を中心に分布すると考えられる。実際に図 4.10, 4.19, 4.20, 4.18 等では、おおよそ 0 の周りに高い分布があるのが伺える。

しかし、図 4.9, 4.12, 4.10, 4.13, 4.15, 4.21, 4.22 等では、明らかに逆相関の分布を示す。また、それとは逆に図 4.11, 4.14, 4.16, 4.17, 4.23 は相関の関係を示す。

これは表情によって、同じ表情を浮かべやすい表情とそうでない表情があるからだと推測できる。例えば逆相関の関係にある AU0 と AU4 を同時に表すのは難しい。(図 4.12) AU0 は上唇の上がり方、AU4 は口角の下がり方なので「上唇を上げながら口角を下げる」「上唇を下げながら口角を上げる」表情は人間にとって難しく、そのためデータの分布に特徴的な値が生じたのだと考えることができる。この様な要因により、AU 情報で構成されるベクトル空間にはデータが存在しない場所とデータが多く存在する場所が点在している。この結果を類似度に活かし、データが集中している部分ではより詳しく、データが集中していない部分ではより広範囲に類似度を計算する手法を取ることも考える事ができるが、今回は採用を見送った。

なお、ヒストグラムの結果より、AU 情報によってばらつきが大きいものとそうでないものがあるということが判明したので、以降の処理は全て標準ユークリッド距離を類似度に採用することにした。

4.2 解析部の評価

得られたデータを元に閲覧ソフトウェアと分析ソフトウェアを動かして実験を行った。実験は以下のように行った。まず閲覧ソフトウェアを利用して、過去のデータの中から基準点を探し出す。次に分析ソフトウェアを利用して基準点と似た表情を持つシーンを選び出す。基準点と似た表情を持つと判定されたシーンは、類似度順にソートされ、1 位から 100 位までの 100 シーンが画面に表示される。そのシーンを以下の 7 パターンに分け、それぞれのシーンがいくつ存在するか比較を行った。

- プログラミング：実際にコードを書いている時である。プログラミングに関する情報をウェブページで検索している時間も含む。(図 4.24)
- サーベイ：論文などを呼んでいる時。また論文をウェブで検索している時を含む。(図 4.25)
- 動画閲覧：ニコニコ動画や映画等を見ている時
- 電子書籍の閲覧：ウェブ上で配信されているウェブ漫画等のコンテンツの閲覧をこう呼称する。視線の動きが特徴的なので、何か変わった表情が取れないかどうか分類した。
- その他（仕事）：電車の予定を調べたり、予約をしたりなどと言った作業
- その他（仕事以外）：仕事以外のページを見ている時等

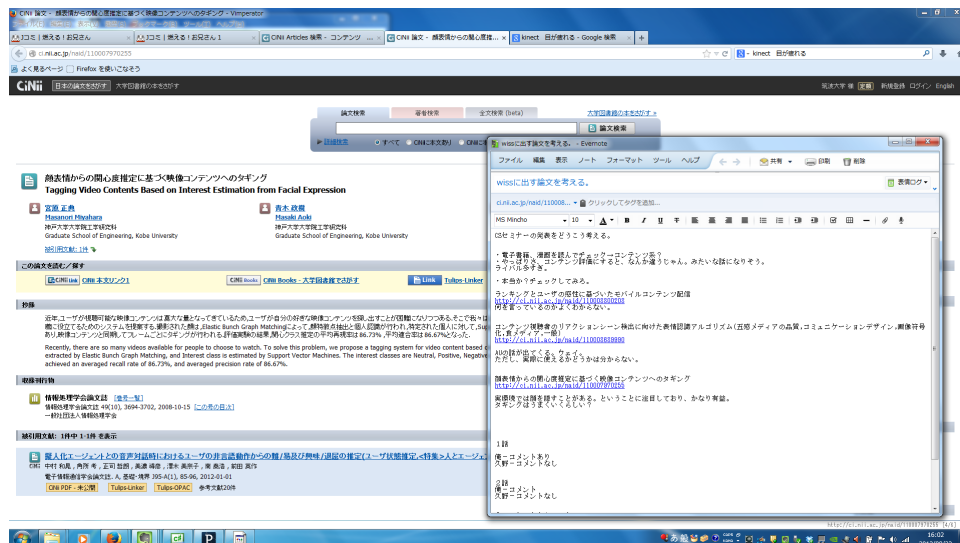


図 4.25: サーベイ時のスクリーンショットの例

- 2013/6/19 1:28 45.560 ～ 2013/6/19 1:35 39.657

を電子書籍に分類する。



図 4.26: 2013/11/18 20:28:29.100 時点の顔

4.2.1 プログラミングのデバッグを行っている時

2013/11/18 20:28:29.100 時点の画像と似ている画像を選び出した。図 4.26 にその時の顔画像を、図 4.27 にスクリーンショットをそれぞれ示す

これは真剣に画面を見てデバッグをしている時の表情データである。デバッグを行っている時と似た表情を持つ顔のシーンを集めることで、過去に行ったプログラミングに関するシーンを検索することが可能なのではと考え、実験を行った。

結果 得られたデータを分類した結果、それぞれ以下の数に該当するシーン数が確認出来た。

- プログラミング：32
- サーベイ：12
- 動画閲覧：30
- 電子書籍：4
- その他（仕事）：12
- その他（仕事以外）：1
- 不明：9

この内容をグラフで表したものを図 4.28 に示す。

プログラミングのシーンに集中して取れると予想していたが、動画閲覧時も同程度の頻度で出現している。しかし、サーベイやその他（仕事）も合わせて作業状態と非作業状態に分類すると

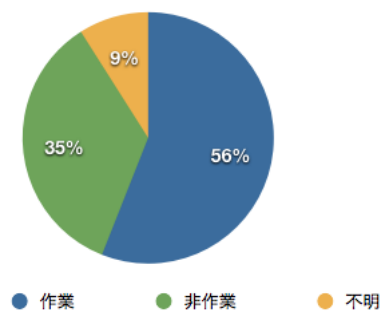


図 4.29: 2013/11/26 18:28:0.324 時点と似ている表情のシーンの作業と非作業分類



図 4.30: 2013/11/26 14:57:39.392 時点の顔画像

4.2.2 動画を見て笑いをこらえている時

2013/11/26 14:57:39.392 時点の顔である．この時点の顔は図 4.30 に示すように，動画を見て笑いをこらえて居る顔である．この顔と似ている顔のシーンを取ることで，筆者が笑いを堪えるようなシーンが見つかるのではないかと考えた．図 4.31 にスクリーンショットを示す．

結果 得られたデータを分類した結果，それぞれ以下の数に該当するシーン数が確認出来た．

- プログラミング：18
- サーベイ：0
- 動画閲覧：60
- 電子書籍：0
- その他（仕事）：6
- その他（仕事以外）：1
- 不明：15

これをグラフにしたものを図 4.32 に示す．
作業状態か非作業状態かの分類を行うと

- 作業状態：24
- 非作業状態：61



図 4.31: 2013/11/26 14:57:39.392 時点のスクリーンショット

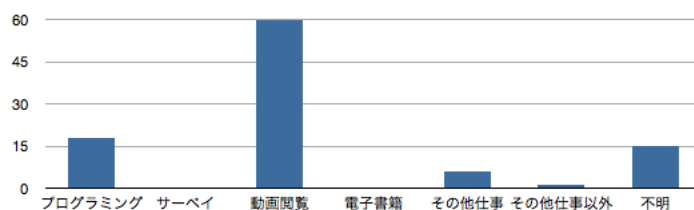


図 4.32: 2013/11/26 14:57:39.392 時点と似ている表情のシーン分類

- その他：15

となり、これを円グラフで表したものを図 4.33 に示す。

動画閲覧が非常に多く、サーベイや電子書籍を見ているシーンは1つも無い。集中して作業を行っている間にはあまり笑いをこらえている様な表情にはならないと考えられる。これは図 4.33 で非作業状態が多く観測されていることから推測することができる。その他（仕事）の時は、旅行の計画を立てていたり、友人と会話をして笑顔になっているシーンが記録されていた。結果として、概ね真面目に作業をしている時でない、笑顔が浮かぶ様なシーンを多く捉えることに成功した。

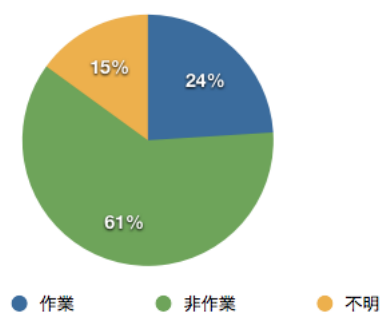


図 4.33: 2013/11/26 14:57:39.392 時点と似ている表情のシーンの作業と非作業分類



図 4.34: 2013/11/26 18:28:0.324 時点の顔画像

4.2.3 顎に手を当てて考えている時の顔

2013/11/26 18:28:0.324 時点の顔と似ている顔を取得した。この顔は、顎に手を当てて何かを考えている時の顔であり、プログラミングの内容について思案している時の顔である。この顔と似ている顔のシーンを取得することで、何か難問にぶつかった時の顔が取得できるのではないかと考えた。

図 4.34 にその時の顔画像を、図 4.35 にスクリーンショットをそれぞれ示す
得られたデータを分類した結果、それぞれ以下の数に該当するシーン数が確認出来た。

- プログラミング：41
- サーベイ：2
- 動画閲覧：41
- 電子書籍：0
- その他（仕事）：7
- その他（仕事以外）：1
- 不明：8

図 4.36 にグラフにしたものを示す。

作業状態か非作業状態かの分類を行うと

- 作業状態：50
- 非作業状態：42

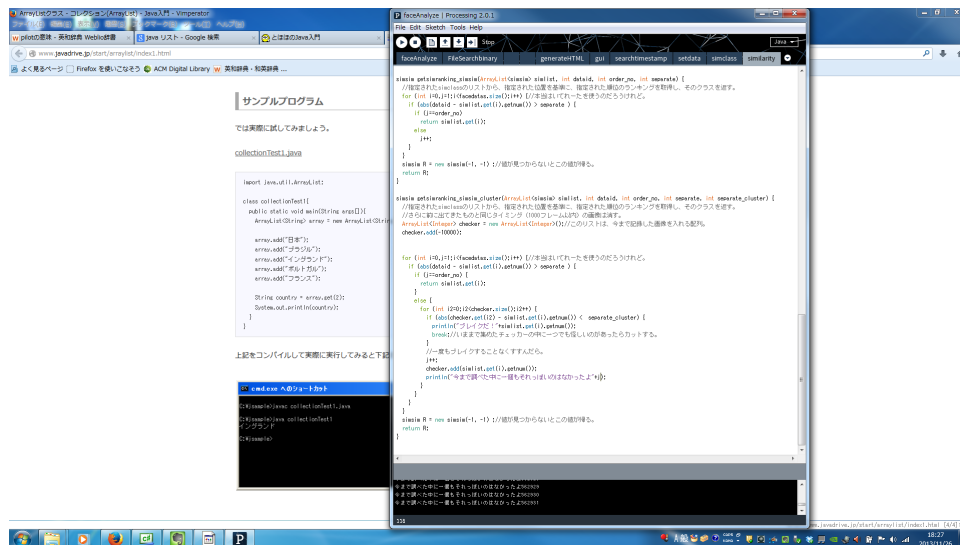


図 4.35: 2013/11/26 18:28:0.324 時点のスクリーンショット

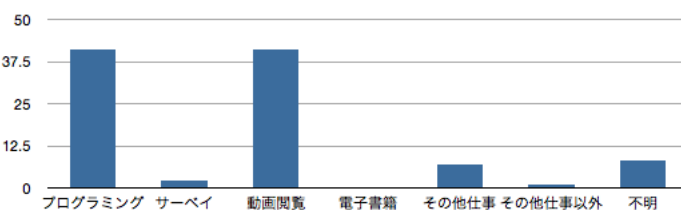


図 4.36: 2013/11/26 18:28:0.324 時点と似ている表情のシーン分類

- その他：8

と言う結果になった．円グラフにしたものを図 4.37 に示す．

図 4.36 に示されている様に，プログラミングと動画閲覧のシーンに大きく二分された結果となった．プログラミング上の問題に当たったシーンだけでなく，動画を見ている間にも何か考え事をしているシーンは存在していた様である．興味深いのがその他の部分であり，作業中にふと気になった謎解きゲーム（ウェブ上で公開されているもの）を解いている時のデータが記録されていた．結果として，考え事をしているシーンの選定はできている様であるが，プログラミングの状態のみを抜きだすことは難しいようだ．

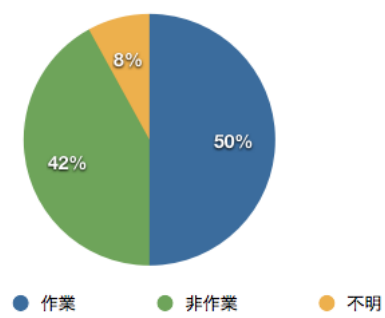


図 4.37: 2013/11/26 18:28:0.324 時点と似ている表情のシーンの作業と非作業分類



図 4.38: 2013/7/6 22:20:21.898 時点の顔画像

4.2.4 動画を見て大きく笑っている時

2013/7/6 22:20:21.898 時点の様に、動画を見て笑っている顔と似ている顔のシーンを取得した。笑いをこらえている顔とはまた別のシーンが取れるのではないかと考えて実験を行った。図 4.38 にその時の顔画像を、図 4.39 にスクリーンショットをそれぞれ示す。

得られたデータを分類した結果、それぞれ以下の数に該当するシーン数が確認出来た。

- プログラミング：11
- サーベイ：0
- 動画閲覧：59
- 電子書籍：0
- その他（仕事）：6
- その他（仕事以外）：11
- 不明：13

これをグラフにしたものを図 4.40 に示す。
作業状態か非作業状態かで分けると

- 作業状態：17
- 非作業状態：70
- その他：13



図 4.39: 2013/7/6 22:20:21.898 時点のスクリーンショット

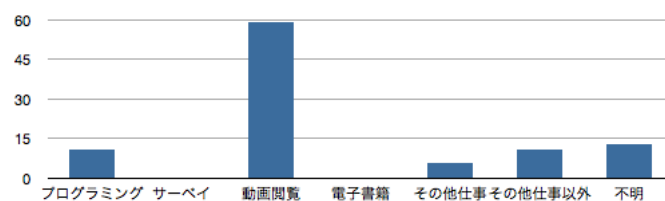


図 4.40: 2013/7/6 22:20:21.898 時点と似ている表情のシーンの分類

という結果になり、図 4.41 に円グラフを示す。

グラフに示されている様に、この結果は非作業状態が圧倒的に多い。特に動画を見ているシーンが多く、動画を見ている時に大きく笑っていることが解る。また作業状態の顔であるが、旅行の飛行機を予約したり、旅行先での計画を立てている時の表情が記録されていた。結果として「笑いをこらえている顔」で取得出来たシーンとよく似ているが、より動画閲覧時に偏ったシーンが取得できた。

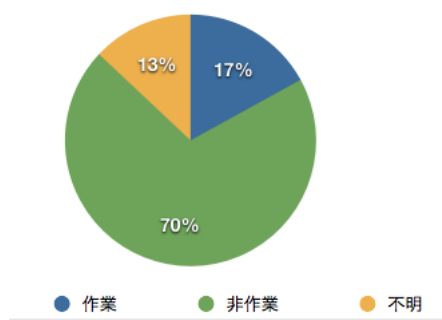


図 4.41: 22013/7/6 22:20:21.898 時点と似ている表情のシーンの作業と非作業分類

第5章 考察

今回取得した4つの状態での表情データを取得してライフログデータからの検索を行った所、作業内容によって異なった結果が取得できることを確認した。特に笑っているような顔の場合は非作業状態が多く観測されることが分かった。画面をじっと見つめている様な状況では、サーベイ等画面を見ているシーンが多く取れた。しかし、動画をじっと見ている様なシーンも取得されたので、作業と非作業の分類は表情によっては難しいということが解る。

笑顔のシーンで分類した結果はかなりの割合で動画に偏っていた。しかし、旅行の計画等、筆者にとって楽しい作業を行っている時のシーンも取得出来ており、過去の情報を思い出すことができるようになった。

本研究ではコンピュータは表情を厳密に認識せず、ただ類似度でデータを取得しているだけであるが、例えばこの表情の時ユーザが「心地よい」と感じているという情報さえあれば、同様に心地よいであろうシーンを探してくることに繋がると思われる。

第6章 結論

ライフログの閲覧と分析に人間の表情を活用するシステムを作成した。このシステムを利用することによって、過去の膨大なライフログデータの中から有益なシーンを選びだすことができるようになった。特に笑顔に関しては本システムを用いることでかなりの精度で検索することができるようになった。表情を活用したライフログ閲覧の有効性についてだが、笑顔だけに限らず、思案顔等の顔でも結果が帰ってくるのが分かった。

今後、このライフログ閲覧手法をインタラクションに応用することによって、「人の顔色を伺うコンピュータ」の実現を目指して行きたい。

謝辞

本論文を完成させる為には多くの方にお力添えを頂きました。指導教官の田中二郎先生からは数多くのご指導を、志築文太郎先生からは、研究の最初から最後まで二年間に渡り丁寧なご指導を頂きました。またチームは異なりますが、三末和男先生、高橋伸先生にも発表等の助言を頂きました。この場を借りて、先生方に感謝申し上げます。IPLABの皆様にも多くのお力添えを頂きました。同期はもとより一足早く卒業していった先輩方からも多くのアドバイスを頂きました。特にゼミで様々な方面から力を貸していただいた IPLAB wave チームの皆様には心より感謝致します。

最後になりますが、ここまで育ててくれた両親，大学生活でお世話になった全ての人に感謝いたします。本当にありがとうございました。

参考文献

- [DS08] Aiden R. Doherty and Alan F. Smeaton. Automatically segmenting lifelog data into events. In *Proceedings of the 2008 Ninth International Workshop on Image Analysis for Multimedia Interactive Services*, WIAMIS '08, pp. 20–23, Washington, DC, USA, 2008. IEEE Computer Society.
- [FTT13] Kurara Fukumoto, Tsutomu Terada, and Masahiko Tsukamoto. A smile/laughter recognition mechanism for smile-based life logging. In *Proceedings of the 4th Augmented Human International Conference*, AH '13, pp. 213–220, New York, NY, USA, 2013. ACM.
- [HHDP12] Javier Hernandez, Mohammed (Ehsan) Hoque, Will Drevo, and Rosalind W. Picard. Mood meter: Counting smiles in the wild. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, UbiComp '12, pp. 301–310, New York, NY, USA, 2012. ACM.
- [JJVS09] Hideo Joho, Joemon M. Jose, Roberto Valenti, and Nicu Sebe. Exploiting facial expressions for affective video summarisation. In *Proceedings of the ACM International Conference on Image and Video Retrieval*, CIVR '09, pp. 31:1–31:8, New York, NY, USA, 2009. ACM.
- [MKK⁺12] Daniel McDuff, Amy Karlson, Ashish Kapoor, Asta Roseway, and Mary Czerwinski. Affectaura: an intelligent system for emotional memory. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pp. 849–858, New York, NY, USA, 2012. ACM.
- [TYA05] Datchakorn Tancharoen, Toshihiko Yamasaki, and Kiyoharu Aizawa. Practical experience recording and indexing of life log video. In *Proceedings of the 2nd ACM workshop on Continuous archival and retrieval of personal experiences*, CARPE '05, pp. 61–66, New York, NY, USA, 2005. ACM.
- [奥浦 11] 奥浦圭一郎, 牛越達也, 河野恭之. Contextual photo browser:写真参与者情報を利用した写真管理システム. 情報処理学会研究報告. HCI, ヒューマンコンピュータインタラクション研究会報告, Vol. 2011, No. 1, pp. 1–7, 2011.

- [岩松 13] 岩松竜也, 岩本健嗣, 松本三千人. 表情とコンテキストを含むソーシャルグラフ構築システムの検討. インタラクション 2013 論文集, インタラクション 2013. 情報処理学会, 2013.
- [牛越 09] 牛越達也, 出射健一郎, 西出亮, 河野恭之. Bluetooth デバイスの検出履歴を用いたユーザ行動の分類. 情報処理学会研究報告. UBI, [ユビキタスコンピューティングシステム], Vol. 2009, No. 2, pp. 1–8, 2009.
- [古川 13] 古川裕士, 濱川礼. 視聴者表情から動画ダイジェストを作成するシステム. インタラクション 2013 論文集, インタラクション 2013. 情報処理学会, 2013.
- [紙谷 97] 紙谷卓之, 山本方子, 丸谷洋二. 瞳位置を利用した困り状態検知のための単位動作認識. 電子情報通信学会技術研究報告, Vol. 96, No. 470, pp. 19–26, 1997.
- [上岡 04] 上岡隆宏, 河村竜幸, 河野恭之, 木戸出正継. I'm here! : 物探しを効率化するウェアラブルシステム. ヒューマンインタフェース学会論文誌, Vol. 6, No. 3, pp. 275–285, 2004.
- [中村 05] 中村和晃, 角所考, 美濃導彦. 学習ログとしての学習者の顔と教材画面の重畳画像の獲得 (先進的学習支援システム). 教育システム情報学会研究報告, Vol. 20, No. 3, pp. 39–44, 2005.
- [平山 07] 平山高嗣, 川嶋宏彰, 西山正紘, 松山隆司. 表情譜 : 顔パーツ間のタイミング構造に基づく表情の記述. ヒューマンインタフェース学会論文誌, Vol. 9, No. 2, pp. 271–281, 2007.