

# 仮想空間と実空間の融合のための 自動色調調整

小野 一義\*    田中 弘美\*    岸野 文郎\*\*

\*立命館大学 理工学部

\*\*ATR 通信システム研究所

仮想現実技術は、臨場感を提供する新しいヒューマンインターフェイス技術として社会的な注目を集めている。仮想現実のリアリティを向上させるための手段として、仮想空間と実空間を融合させることが有望視されているが、融合領域の抽出法や融合の不自然さの解消など解決すべき課題が多い。そこで本稿では、対象の3次元形状とテクスチャー変化を観測した画像の分析を基に、対象の3次元CGモデルと実写ビデオライブ映像の融合境界を抽出し、両者をリアルタイムで融合する方法を提案する。さらに、自然で滑らかな融合を実現するために融合境界線に沿った色調の差異を自動調整する方法を示す。本手法を実時間顔表情生成に応用した実験結果を示す。

## Adapting Color Properties for Smooth Merging Virtual and Real Spaces

Kazuyoshi ONO\*    Hiromi T. TANAKA\*    Fumio KISHINO\*\*

\*Faculty of Science and Engineering, Ritsumeikan University

\*\*ATR Communication Systems Research Laboratories

Virtual reality technology provides a new methodology for human interface with realistic sensation. We describe the method of adapting color properties for smooth merging virtual and real spaces. We extract regions suitable for merging virtual and real spaces in order to enhance reality based on analysis of both object's 3-D shape reformation and its surface texture change. Then we adapt color properties to decrease the contrast at the boundary of the merged region for smooth merging on realtime. We apply this technique to realtime animation of realistic facial expression.

## 1 はじめに

3次元仮想空間 (VR 空間) 技術は、臨場感を提供する新しいヒューマンインターフェイス技術として社会的に大きな注目を集めている。リアリティを与える技術の中心は、コンピュータグラフィックス (CG) の描画機能を用いて視覚的に提示する方法である。しかし VR 空間のすべての構成物体の幾何学的構造や物理的拘束などを CG モデル化することは不可能であるため、VR 空間の与えるリアリティを向上させるため手法として、仮想空間と実空間を融合 (CG と実写を融合) させることが有望視されている。

CG と実写映像の合成は、映像製作分野では盛んに行われてきたが、CG と実写の区別がつかないほど自然に合成させるには多大な人手を要しており、実写を投影する領域 (以下、実写領域) の境界を自動的に抽出する方法の開発が望まれている。

また最近の VR 研究分野では、実写映像を用いてリアリティの高い VR 空間を生成する研究が活発化してきている。遠隔制御の操作性等の向上を図るために、背景の実写ビデオ映像に対象の CG ワイヤフレームモデルを重畳表示する Augmented Reality[5-7] の提案や、CG モデル化の煩雑さを軽減するために、背景を複数の実写ステレオ画像を用いて立体視し VR 空間に融合させる方法 [13] があげられる。

しかし現状では、実写画像を用いた場合は任意方向からの立体視が困難であることや、動的な対象に対して正確に実写を投影することが困難であることなどから、適用範囲は「静的な」背景に限られている。さらに融合に有効な実写領域の自動抽出や、融合境界が不自然さに際立って見えることなど解決すべき課題が多い。

そこで本稿では、対象の3次元 CG モデル上に実写ライブビデオ映像を投影することにより、実写映像の適用を「動的な」対象に拡張する仮想空間と実空間の実時間融合法を提案する。本手法は、従来の VR 空間における、複雑で微妙な動き変化生成のための繁雑な3次元 CG モデル化から解放

し、かつ任意方向からの立体視を可能にするものである。

まず、対象の3次元形状とカラーテキストチャー変化を観測した距離画像とカラービデオ画像の分析を基に実写映像の有効性を評価し、仮想空間と実空間の融合境界を抽出する。

次に、撮影条件の違いに起因しかつ融合の不自然さを知覚させる主因をなる CG モデルと実写領域の色調の差異をカラヒストグラム分析を用いて自動調整することにより、対象のライブ映像を実時間で反映する自然で滑らかな融合を実現する。

本手法をリアリスティックな実時間表情生成に適用し、人間が鋭敏に知覚する人物の顔を対象に評価を行った。

## 2 仮想空間と実空間の融合

### 2.1 実時間融合の問題点

リアリティの高い仮想空間を生成するためには、空間を構成する各対象の3次元形状や表面テキストチャー、またそれらの複雑で微妙な変形や運動を高精度にモデリングすることが必要である。

観察画像を基に実時間で対象の動きをモデリングするためには、画像理解 (CV) の立場から以下の困難な問題の解決が必要である。

1. 実時間で特徴点を抽出する。
2. 実時間で特徴点を追跡する。
3. 実時間で特徴点を対応付ける。
4. 実時間で特徴点の動きパラメータを抽出する。

また、モデリングされた対象の動きを仮想空間に生成し提示するためには、CG の立場から以下の困難な問題の解決が必要である。

1. 表示・生成のための動きパラメータ設定。
2. 実時間動き生成のための変形計算コスト。

つまり、よりリアルで高精細な動きを生成するためには、より多くの特徴点とパラメータを設定しなければならず、膨大な計算量が必要である。

動きが複雑でしかも動きの数が多い対象を、実時間で仮想空間に反映させるのは、コンピュータの計算力が飛躍的に増大し、CG画像の写実性(フォトリアリティ)が向上した現在においても、ここで述べた実時間処理の制限のため表現力に限界がある。

そこで、VR空間の与えるリアリティを向上させるため手法として、仮想空間と実空間を融合(CGと実写を融合)させることが有望視されている。

## 2.2 融合方法

仮想空間と実空間の融合に関する研究としては、背景に実写を用いる方法と、対象物体にである実写を用いる方法に分類することができる。

### 1. 背景に実写を用いる方法

- (a) 実世界をビデオカメラ等で立体視している系に、ワイヤフレーム等を重畳表示するアプローチ。遠隔操作性の向上をはかろうとする Augmented Reality [5, 6, 7] 等の研究例がある。
- (b) シースルー型 head-mounted display (HMD) を用いて、ディスプレイ上に実空間に関連したCGを写すことにより実空間と仮想空間を融合するアプローチ。超音波を使って得られた3次元画像をシースルー型 HMD に映すことで、実空間との融合を図った、M.Bajura, H.Fuchs 等 [3] の研究例がある。
- (c) ステレオ画像を背景としてCGに融合し、立体視するアプローチ。田村、北村 [13] 等の、HoloMedia システムなどの研究例がある。これは、複数のカメラから得られたステレオ画像をEPI(Epipolar Plane Image) を用いて補間することで、立体視できる系を提示する。

また田村 [13] では、従来のCGモデル系と実写を使った非モデル系のアプローチの特徴を以下に論じている。モデル系は、3次元空間

内の物体の幾何形状と表面テクスチャ情報について、その構造と属性を明示的に示した系であり、非モデル系は、3次元の情報を持たずに、形状モデリングが難しい対象に対して実写を活用して立体映像を生成・提示できる系である。モデル系の特徴は、一旦モデル化し仮想空間で表現できた対象については、任意の視点からの立体映像が生成できるが、対象の変化に追従した画像を実時間で生成するのは困難である。これに対し、非モデル系の特徴は、モデリングコストが少ないが、限られた範囲の視点からの立体視に限定される。

2. 対象物体に実写を用いる方法 対象物体に実写を用いる方法としては、仮想空間にモデリングの難しい人物像等をクロマキー技術を用いてリアルタイムでCG画像に合成するのが一般的であり、TVや映画製作によく使われている。しかし、対象の実写映像を取る場合、容易に二値化できるような色・模様を背景としなければならないなどの撮影条件に制約があり、また対象を2次元の情報として融合するため、受け手の視点に追従した仮想空間を生成・提示できない。

## 3 実写有効領域

### 3.1 画像分析による有効性の評価

対象の動きを再現する際に、仮想空間の固定した3次元CGモデルに実空間のビデオライブ映像をテクスチャとしてマッピングし、リアルな仮想空間を生成する。

実写映像の有効領域を距離画像とカラー画像を用いて、テクスチャと3次元形状の分析を基に抽出する手法を示す。テクスチャと3次元形状を静的特徴と動的特徴についてそれぞれ分析する。

#### 3.1.1 テクスチャ分析

##### ● 静的特徴

テクスチャの複雑を解析するために、色数・

一様性・方向性・規則性・密度などについて分析する。

- 動的特徴

テクスチャは対象の動きによって変化するが、その変化量を色差や、オプティカルフローなどを用いた移動量を計測し分析する。

本研究では、対象の動きに対する色差を求める。色差を評価する色差式には、Godlove 色差を用いた。これは、輪郭を持った色票について定義されたものである。つまり、均等色空間における2色間の色差と人間が感じる色差との対応関係が利用されていることを利用することにより、人間の感覚に近い解析が可能である。比較する2色のHSV値を $(H_1, S_1, V_1)$ 、 $(H_2, S_2, V_2)$ とすると、Godlove 色差は次式で定義される。

$$\Delta E(\text{Godlove}) = \{(2S_1 S_2)\{1 - \cos(2\pi \cdot \Delta H/100)\} + (\Delta S)^2 + (4\Delta V)^2\}^{1/2} \quad (1)$$

$$\text{但し、} \Delta H = |H_1 - H_2| \quad (2)$$

$$\Delta S = |S_1 - S_2| \quad (3)$$

$$\Delta V = |V_1 - V_2| \quad (4)$$

### 3.1.2 3次元形状分析

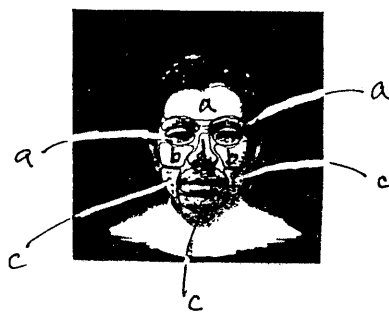
- 静的特徴

対象の3次元形状の複雑さを分析する。つまり対象の表面が、平面や2次元曲面などの単純な形状をしているか、曲率が複雑に変化する自由曲面であるかを分析する。

- 動的特徴

対象の動きに応じて3次元形状は変化する、このときの動きの数・方向・大きさ・奥行き方向の形状変化を分析する。

顔表面を対象とした場合は、実写有効領域の定義に従って、図1に示すように、顔表面は領域分割される。図1の領域(a)が、実写有効領域の候補として選択できる。形状変化については本来、自ら測定し分析することが正であるが、今回は上野 [14] に基づき考察した。



- (a). 皮膚表面での2次元運動をする領域
- (b). 皮膚表面に対して膨らみや沈みをする領域
- (c). 3次元運動を柔軟にする領域
- (d). ほとんど運動しない領域

図1: 分析にもとづく顔表面の領域分割 (上野 [14] p15: 図8より抜粋)

### 3.2 実写有効領域

固定の3次元CGモデルにビデオライブ映像をする場合、3次元形状がダイナミックに変形する領域、あるいは、モデルの変形が容易であるような領域に対しては、実写映像は有効ではない。これに対し、3次元形状変化が小さく対象の表面上を滑るような動きをする領域に対しては、3次元の形状を変化させなくとも実写を用いてリアルに再現できる。

このような実写を融合することによりリアルな仮想空間が得られる領域を実写有効領域として抽出する。

つまり、領域内の内側にテクスチャ変化のほうに形状変化より複雑な領域を内包し、その周囲に形状とテクスチャ変化とも小さい境界領域に囲まれる領域を実写有効領域とする。

## 4 主色分析による色調調整

前章で抽出した領域に実写を融合する際、撮影照明条件の違いによる仮想空間の3次元モデルと実写ビデオ映像の色調の差異のため、融合境界が際立ってしまう。

この色調の差異を緩和し、かつ実空間のライブ

情報を反映させるために、CGモデルの色調をライブビデオ映像の色調に近づける。これにより、スムーズな融合をはかり、よりライブの対象に近い色調を仮想空間内に生成する。

#### 4.1 主色の抽出と分析

色調の差異を測定するために、実写有効領域内の量的かつ、質的に支配的な色を主色として抽出する。

まず、カラーヒストグラムを用いて量的な主色候補を選出し、つぎに、境界線近傍の色分布から、融合のなめらかさを決定する質的に支配的な色を主色として抽出する。

主色分析には、各画像のR(赤)、G(緑)、B(青)をマンセル表色系で用いられる色相(Hue)、彩度(Saturation)、明度(Value)に変換し分析した。主色は、色相の範囲を  $min < Hue < max$  指定する。

色調調整のアルゴリズムを次に示す。

1. 2つの入力画像をそれぞれHSV変換する。
2. 主色領域 ( $min < Hue < max$ ) のHSVそれぞれのヒストグラムを作る。
3. Hのピーク値 ( $H_{peak}$ ) と、 $S \cdot V$ の平均値 ( $Save, Vave$ ) をそれぞれ求める。
4. 次の変換式でCGモデルのH、S、Vをそれぞれ変換する。

$$H_{CG} = H_{CG} + \Delta H \quad (5)$$

$$\Delta H = H_{peak\_video} - H_{peak\_CG} \quad (6)$$

$$S_{CG} = S_{CG} \cdot \Delta S \quad (7)$$

$$\Delta S = S_{peak\_video} / S_{peak\_CG} \quad (8)$$

$$V_{CG} = V_{CG} \cdot \Delta V \quad (9)$$

$$\Delta V = V_{peak\_video} / V_{peak\_CG} \quad (10)$$

5. HSV逆変換をする。

## 5 仮想空間と実空間の実時間融合法

以下に全体的な処理の流れを示す。なお、入力には、対象の3次元CGモデルと実写ライブビデオ映像である。

step1: 実写有効領域の抽出

step2: 実写有効領域内の主色分析と色調調整

step3: 3次元CGモデルとビデオ映像の融合

step4: 実写融合CGモデルの実時間出力

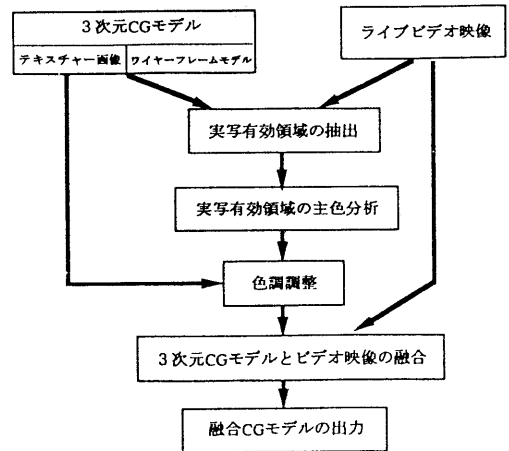


図 2: 処理の概要

対象の任意の動きに対して固定の領域の実写映像を得るために、図3に示すように対象に固定したカメラからのビデオ映像を使うことで実時間で特徴対応問題を解くことを回避している。また、CV・CG両分野の問題点を回避することができる。本

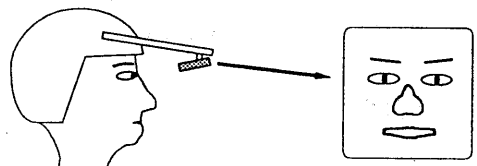


図 3: Head Mounted Camera

研究の初期実験として、色差のみを対象に、テクスチャ変化を分析した。

顔の表情をあらわすのに重要ないくつかの部分

のうち、目の部分が他の部分に比べ、表情の変化に対して形状変化が少なく、その大部分は平面的な変化と考えられることから、図 5 に示す目の周りの眼鏡形状領域を実写有効領域とした。



図 4: テクスチャ変化



図 5: 実写有効領域

2.2 節の方法はいずれも融合する仮想空間と実空間が距離的に異なり、距離的に近接している場合や複雑に絡みあったものを融合のは簡単ではない。よって、融合境界と奥行き不連続が一致するため融合境界が比較的に際立たない。しかし、本研究では、融合境界で奥行きは連続しており、色調の差異により融合境界が際立ってしまう。

そこで色調の違いを自動調整して融合境界の際立ちを緩和し、スムーズな融合をはかる。

## 6 実験

3次元CGモデルはオフラインで、Cyberware 3D Digitizer を用いて対象を撮影した。同 Digitizer から入力される距離情報は円筒座標系で表現されており、Utility Software により、3次元ワイヤフレームモデルに変換される。ビデオ映像は、対象の人物がかぶったヘルメットに取り付けた CCD

カメラ(図 3) から得た。CCD カメラは Panasonic の WV-KS102 を用いた。解像度は 628(H)x492(V) である。撮影したビデオ映像をシリコングラフィックス社のワークステーションに、同社のビデオキャプチャボード VideoLab によってその映像を実時間で取り込んだ。

実写有効領域とした領域に対し主色分析をした。ヒストグラムから、主色を肌色領域 ( $0 < Hue < 30$ ) として抽出した。3章で示したアルゴリズムに従い、3次元CGモデル(図 6)の色調を調整した。

表 1 に、2つ入力画像の画像の主色分析による色調調整の結果を示す。また、ビデオ映像と3次元CGモデルの2つの入力画像の H、S、V 値のヒストグラムとビデオ映像と色調補正後の3次元CGモデルの H、S、V 値のヒストグラムを図 8 に示す。

表 1: 色調調整結果

	H peak	S ave	V ave
Video Image	15.00	0.344	0.345
3D CG Model	10.00	0.579	0.775
Result	10.00	0.582	0.710

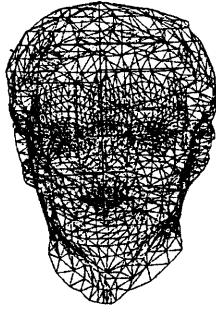
本実験では、手動で融合領域をビデオ画像から指定し融合する。色調調整後の3次元CGモデルとビデオ映像の融合結果を図 7 に示す。



図 7: 融合結果

## 7 おわりに

形状及びテクスチャー分析に基づいて仮想空間と実空間を融合することを提案した。また、融合



(a). 3次元CGモデル

(b). 実写ビデオ映像

図6: 入力画像

境界領域の色調の差異の補正方法について提案した。本手法をリアルタイム顔表情生成に適用した。この手法の利点は実写有効領域に対しては、3次元形状を変化させなくともリアルな対象の生成ができること。また、ライブの実空間の色調を忠実に再現できることである。

今後、さらにテクスチャと3次元形状の分析を進め、融合領域の抽出を自動化したい。また、融合境界を目立たせる色調以外の要素について考察していきたい。

## 謝辞

貴重なデータを貸して頂く等、御協力下さった(株)ATR通信システム研究所の越知 武氏、エム・アイ・システムズ(株)の池田 孝利氏に、感謝の意を禁じ得ません。

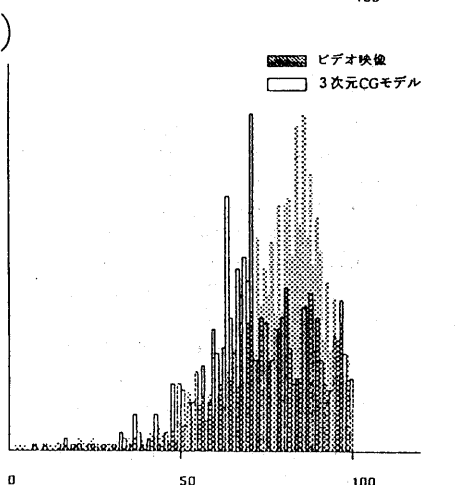
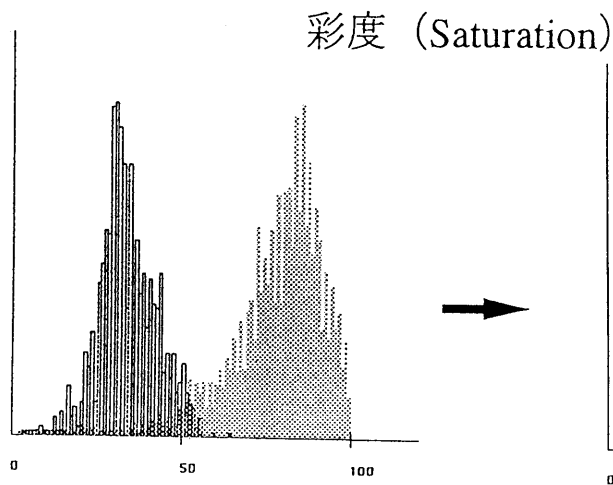
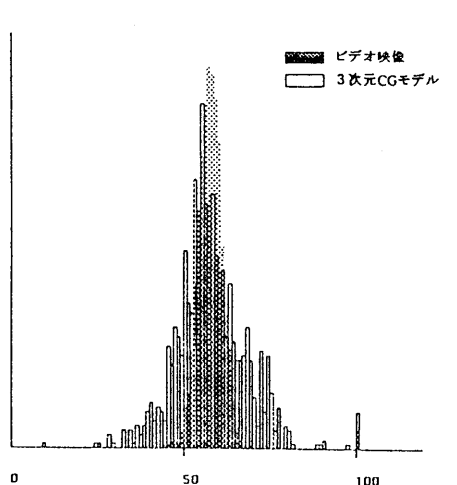
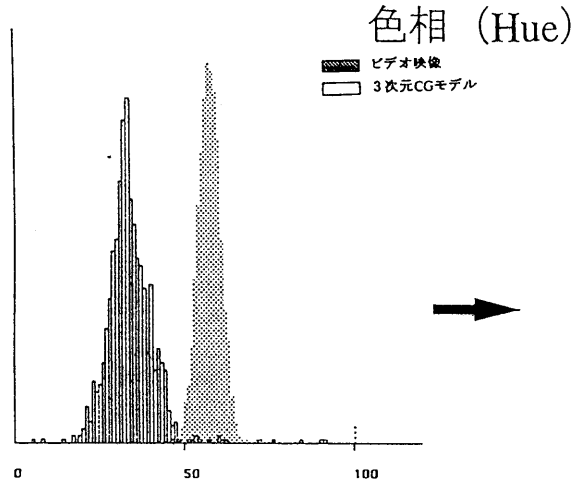
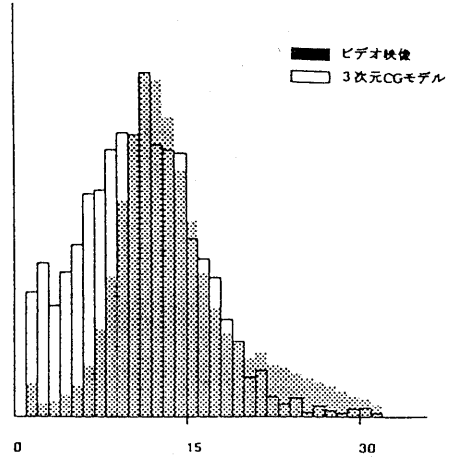
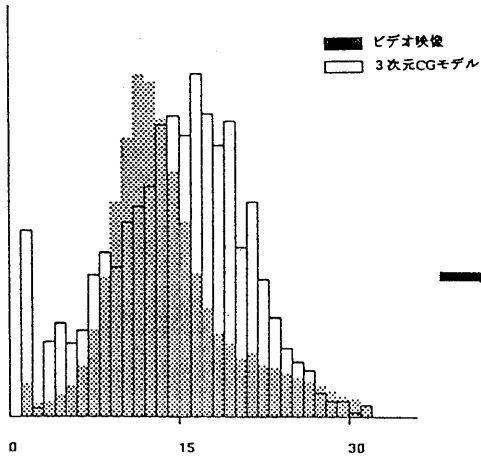
## 参考文献

- [1] H.D.Foley,A.Van Dam:“Fundamentals of Interactive Computer Graphics”, Addison-Wesley Publishing Company(1982).
- [2] 岸野, 山下:“臨場感通信のテレコンファレンスへの適用”, 信学技法 IE89-35(1989.9)
- [3] M.Bajura,H.Fuchs:“Merging Virtual Objects with the Real World”,SIGGRAPH,Computer Graphics,pp203-210(July,1992)
- [4] 大谷 淳:“人物像の認識と合成: 実時間処理を目指して”, ATR ジャーナル, No.13, pp10-13(1993 春)
- [5] P.Milgram:“Enhancement of 3-D Video Displays by Means of Superimposed Stereo-Graphics”, Proceedings of the Human Factors Society 35th Annual Meeting(1991)
- [6] P.Milgram:“Applications of Augmented Reality of Human-Robot Communication”, Proceedings of the IEEE/RSJ International Conference(1993)

- [7] P.Milgram:“Mixed reality: Research issues in merging real and virtual worlds”, ATR Workshop on Virtual Space Teleconferencing(1993).
- [8] P.Milgram:“A class of displays on the reality-virtuality continuum” SPIE Vol.2351 Telemanipulator and Telepresence Technologies(1994)
- [9] P.Milgram:“A Taxonomy of Mixed Reality Visual Display”, IEICE TRANS.INF. &SYST., Vol.E77-D,No,12 pp1321-1329(1994)
- [10] R.Nevatia:“画像認識と画像理解: 視覚的マシン知覚”, 啓学出版株式会社, 東京 (1986).
- [11] 坂口, 森島, 大谷, 岸野:“3次元計測に基づく顔表情変化の分析と合成”, 信学技報, pp 61-68(1994-01)
- [12] 高木 幹雄, 下田 陽久:“画像解析ハンドブック”, 東京大学出版会, 東京 (1991)
- [13] 田村 秀行, 北村 素子:“3次元仮想空間生成・表示へのモデル系&非モデル系アプローチ”, Human Interface Vol.9, pp.279-286(1994).
- [14] 上野, 小野, 森島, 原島:“自然な表情合成のための頭部高精細ワイヤフレームの構成とその階層的制御について”, 信学技報, pp 9-16(1992-12)

色調調整前

色調調整後



明度 (Value)

図 8: ヒストグラム