

## 画像検索を用いたモノクロ画像のカラリゼーション

森田 千晶<sup>†</sup> 増永良文<sup>‡</sup>

<sup>†</sup>お茶の水女子大学大学院 人間文化研究科博士前期課程 数理・情報科学専攻

<sup>‡</sup>お茶の水女子大学 理学部情報科学科

〒112-8610 東京都文京区大塚 2-1-1

E-mail: <sup>†</sup>chiaki@db.is.ocha.ac.jp, <sup>‡</sup>masunaga@is.ocha.ac.jp

**あらまし** 近年,古い資料や写真のデジタル化などに伴い,モノクロ画像の自動着色技術の研究がすすめられている.そのひとつとして,色合いの参考となる画像(ソース画像)を用いてその色合いをモノクロ画像に転送することによってカラリゼーションを行う手法が提案されている.この手法はユーザの作業負担を軽減し自動的なカラリゼーションを実現させるが,そのために必要となる“ユーザの主観にあうソース画像”の用意が困難である.そこで本研究では,この手法を前提とし,“ユーザの主観にあうソース画像”を,画像検索技術を組み合わせることによってウェブ上から取得してくる方法を考案する.また,本稿ではこのシステムのための予備実験として,ウェブ画像検索の検索キーワードの選択方法に関する考察をおこなう.

**キーワード** ウェブ画像検索, 画像類似検索, カラリゼーション, 主観

## Colorization of Grayscale Images by Image Retrieval

Chiaki MORITA<sup>†</sup> Yoshifumi MASUNAGA<sup>‡</sup>

<sup>†</sup> Graduate Division of Mathematics and Computer Science (Master's Program),  
Ochanomizu University

<sup>‡</sup> Department of Information Science, Faculty of Science, Ochanomizu University  
2-1-1 Otsuka, Bunkyo-ku, Tokyo, 112-8610 Japan

E-mail: <sup>†</sup>chiaki@db.is.ocha.ac.jp, <sup>‡</sup>masunaga@is.ocha.ac.jp

**Abstract** Recently, the research on automatic coloring technologies for grayscale images is extensively done due to the needs for digitalizing old materials and photographs. There is a technique for colorizing grayscale images by transferring color from a source color image to a grayscale image which is the destination. This technology seems to be attractive, but there is a problem of how to provide a source image that reflects user's subjectivity. In order to resolve this problem, we introduce a method of acquiring the source image along user's subjectivity from the Web by the image retrieval. A preliminary experiment is done in order to identify a set of keywords for effective retrieval of images from the Web.

**Keyword** Web Image Retrieval, Similarity Image Retrieval, Colorization, Subjectivity

---

<sup>†</sup>お茶の水女子大学大学院 人間文化研究科博士前期課程 数理・情報科学専攻  
Graduate Division of Mathematics and Computer Science (Master's Program), Ochanomizu University

<sup>‡</sup>お茶の水女子大学 理学部情報科学科  
Department of Information Science, Faculty of Science, Ochanomizu University

## 1.背景と目的

カラリゼーションとは、モノクロの画像や動画などに対して、コンピュータ技術を用いて色彩をつける技法である。この技法は画像を扱う様々な分野での応用が考えられる。近年、古い写真のデジタル化に基づく修復作業や白黒映画のリバイバル化が進められ、カラリゼーションによる色の復元が望まれている。また、3Dコンピュータグラフィックス画像や、衛星データ・CTRデータからの可視化画像などの作成においても着色作業は大きなウェイトを占めるため、カラリゼーションの需要が見込まれる。

カラリゼーションは、市販のフォトタッチソフトを用いても行うことができる。一般的には、ユーザは最初に手作業で画像のパーツ部分の領域分解を行う。そして、各領域に色を重ねたり、色調変換を行ったりすることにより着色を施していく。しかし、この方法では膨大な時間と労力を要することになり、非常に効率が悪く、動画のカラリゼーションにいたっては、フレームごとの着色作業となるため、この何千倍もの作業時間が必要になり、製作の負担が非常に大きい。また、各輝度値には対応する色が複数存在するため、このような直接的に色を選択する方法では自然な色の選択が難しい。近年、新しいカラリゼーションの手法として、色合いの参考となる画像(以下、ソース画像という)を用意し、その色合いを転送することによって着色を行う研究がなされている。この手法では、着色のほとんどの過程を自動化することができるためユーザの負担を大幅に減らすことができる。しかし、自分のイメージに沿った色合いをもち、さらにカラリゼーションを行うために有効であるソース画像を用意することは非常に困難となってくる。

そこで本研究では、前述のカラリゼーション手法の発展として、ソース画像を取得するための画像検索手法を考案する。画像取得場所は、膨大な量の画像を保持

するウェブ上の画像とする。これにより、着色までの全過程を自動化し、主観に忠実に自然な色味での着色を施すことができるカラリゼーションシステムを提案する。

## 2.関連研究

モノクロ画像のカラリゼーションにおける研究はさまざまな方向でのアプローチがなされているが、これらは大きく2種類に分類ができる。一つは、ユーザがカラーパレットからの色を選択し、割り当てていく方法である。ポイントやラインなどでユーザが最小限の色指定を行うことにより、自動的に適正範囲に着色を施す研究などがなされている[1][2]。この手法はユーザ側に着色自由度が非常に高く、作業負担も軽くなっているが、色の正しい選択という部分に関しては考慮がなされていない。もう一つは、色味の参考となるソース画像を用意し、その画像と対象画像の輝度やテクスチャのマッチングをみることにより、色味を転送する方法である[3][4]。この手法は、実際に存在する色味を使用するので不自然な色合いにはなりにくい。しかし、ソースとなるべき画像が手元に存在していないと着色が行えないという難点もあり、またそのような適正ソース画像の取得も難しい。本研究では特に後者の手法に着目し、画像検索技術を用いてソース画像取得の問題を解決する。

画像検索技術としては、Google イメージ検索[5][6]のようなキーワード検索を応用したウェブ画像検索技術や、画像の特徴量やメタデータを元に画像DBから検索を行う画像類似検索技術などがある。画像類似検索の特徴量には、色、テクスチャ、形状、位置情報などが用いられ、これらの特徴抽出にはウェーブレット変換が使われることが多い。さらに、これらの特徴量からメタデータ自動抽出し検索を行う研究も進められている[7]。また、

goo ラボ[8]では、従来のウェブ画像検索技術と画像類似検索技術を組み合わせることにより、画像や動画の検索効率向上を目指した新しいマルチメディア検索システムの公開実験も行っている。

### 3. カラリゼーション

本システムでは、カラリゼーション方法として、welsh ら[3]による、ソース画像を用いて色転送をおこなう手法を採用する。この手法では、ユーザは入力としてカラリゼーションの対象となるモノクロ画像と、色味の参考となるソース画像を用意する。そして、モノクロ画像内の各ピクセルに対して、ソース画像からサンプリングしたサンプルの中で最もマッチングするものを選択する。マッチングはピクセル輝度と近傍統計の加重平均を用いることによって決定される。このときのソース画像サンプルの色値をモノクロ画像の各ピクセルに転送することによってカラリゼーションを行っている。

処理をおこなうための色空間には、色度図が均等である直交三次元座標系から構成される Lab 色空間を用いる。Lab 色空間は、それぞれ L: 無色輝度チャンネル, a: 黄-青チャンネル, b: 赤-緑チャンネルの情報をもっており、チャンネル同士で互いに影響を及ぼしにくいという特徴をもっている。これは、輝度を固定したまま色彩を調整することができるため、カラリゼーションにおける色空間としても多く用いられている。RGB 値を Lab 値に変換するには以下の式を用いる [4]。

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.3811 & 0.5783 & 0.0402 \\ 0.1967 & 0.7244 & 0.0782 \\ 0.0241 & 0.1288 & 0.8444 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

RGB 値を一度 LMS 値へ変換する。この段階で、色空間のデータにゆがみが生じているため、ログ空間へと変換することで修正する。

$$L = \log L$$

$$M = \log M$$

$$S = \log S$$

それから以下の式を用いて Lab 色空間へと変換する。

$$\begin{bmatrix} 1 \\ a \\ b \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{6}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -2 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix}$$

### 4. システム概要

本研究では、最低限のユーザ入力で、より自然な着色を可能とするような、モノクロ画像の自動カラリゼーションのシステムを提案している。システム全体は、ソース画像とのマッチングを利用して着色を施す“カラリゼーション”と、そのソース画像をウェブ上から取得してくる“画像取得”の大きく二つの機能から構成されている。本稿では、特に後者の“画像取得”部分に焦点をあてる。“画像取得”は、キーワードによりウェブ上の画像からユーザの主観に沿った画像を複数取得してくる“ウェブ画像検索フェーズ”と、それらの画像の画像データベースを用い、画像特徴量を考慮することによって最も適当となるソース画像を決定する“画像データベース検索フェーズ”に分かれる。システム全体の概要を図1に示す。

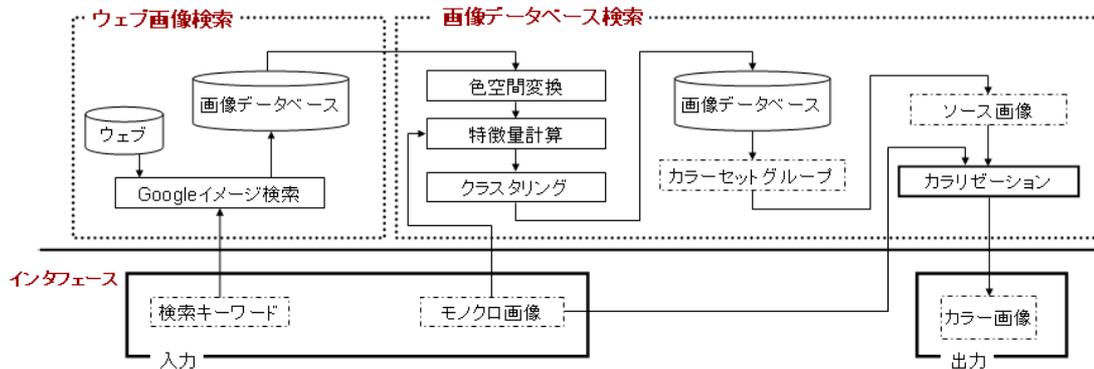


図 1. カラリゼーションシステムの概要

#### 4.1. システムの特徴

カラリゼーションでの着色は、対象物がもつ正確な色合いを再現するためのカラリゼーションと、ユーザがもつ“色合いのイメージ”（以下、主観という）を元に着色をおこなうカラリゼーション、という二つの方向性に大別される。本システムでは、後者に視点を向けて、ユーザの主観をカラリゼーションに反映させることに重点をおく。これにより、芸術的用途での利用まで含めた幅広い用途での応用に対応できるシステムを目指す。

##### 4.1.1. カラーセット

本システムでは、画像がもつ“色合い”を Lab 色空間での各ピクセルのもつ Lab 値の集合体として定義する。この集合体をカラーセットとよび、色合いを判断するための基準とする。

##### 4.1.2. 主観の反映

ユーザは、本システムにより複数表示される候補画像の中から画像を選択することで、ユーザのもつ主観とカラーセットを結びつける。この結び付けられたカラーセットを主観カラーセットとよぶ（図 2）。これを元に、類似カラーセット

グループ（同章 3 節）内から最もカラリゼーションに適する画像を正式なソース画像として選択する（図 3）。これにより、ユーザの主観を反映しながら精度の高いカラリゼーションを実現させる。また、ウェブ画像検索の検索ワード選択においても、より主観を反映させるような考察をおこなっている。詳細に関しては、5 章にて後述する。

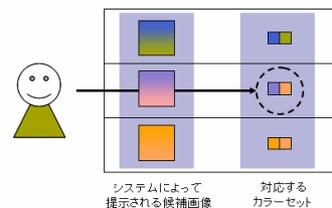


図 2. 主観カラーセット定義の概念

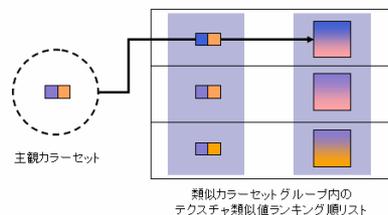


図 3. 主観カラーセットからソース画像決定までの概念

## 4.2. ウェブ画像検索

ウェブ画像検索フェーズでは、キーワードを用いてウェブ上の膨大な量の画像からユーザの主観に沿った画像を検索し、それを用いてソース画像候補の画像データベースを構築する。本システムでは、Google イメージ検索の利用を考える。Google イメージ検索では、検索ワードの選択が非常に重要となる。本稿では、主観を反映させた結果を得るために、5章において予備実験として検索ワードに関する考察をおこなった。

### 【Google イメージ検索】

Google イメージ検索は、大手検索エンジンである Google が提供するサービスであり、ウェブ上の画像を任意のキーワードを用いて検索する。ウェブ上のあらゆるコンテンツに存在する4億以上の画像インデックスを保持しており、独自のアルゴリズムで高品質の検索結果を提供している。索引化の対象になっているものは、画像が含まれるページ内の隣接テキスト、img タグの alt オプション内のテキスト、ファイル名、リンク元アンカーテキストなどが考えられる。検索結果のランキングには、PageRank とよばれるアルゴリズムが用いられている。これは、ウェブページの重要性を他のウェブページからのハイパーリンクの数で評価する。アルゴリズムは以下の計算式で簡単に表すことができる[6]。

$$PR(A) = (1-d) + d \left( \frac{PR(T1)}{C(T1)} + \dots + \frac{PR(Tn)}{C(Tn)} \right)$$

PR(A)はページ A の PageRank である。C(T1)はページ T1 から外に向けられるリンクの数を表す。d は制動係数で、 $0 < d < 1$  の範囲をとる。Google では、通常 0.85 に設定されている。つまり、ページ A の PageRank は、ページ A に向かってリンクしている(インリンク)ページの PageRank を、それぞれのページの外向きリンク(アウトリンク)で割ったものの総和になる。Google の検索結果のランキングは、この PageRank アルゴリズムとその他多くの

要因の重み付けによって、決定がなされていると思われる。

また、Google イメージ検索では、様々な検索オプション提供されている。本システムでは、これらのオプションを用い、画像形式を JPEG のカラー画像に絞った検索をおこなう。これは、色数を多く表現できる JPEG は写真に、GIF や PNG は図やマップなどに用いられることが多いからである。カラリゼーションのソース画像となる得る画像は、写真のように色数の多いカラー画像が有効であると考えられるため、以上のように対象の形式を指定することで検索精度をあげる。

## 4.3. 画像データベース検索

画像データベース検索フェーズでは、ウェブ画像検索フェーズで作成した画像データベースを用いて、ユーザの主観を反映したソース画像の取得をおこなう。

従来の画像類似検索システムが、色、テキスト、形状、位置情報などを用いて類似度を計算していくのに対して、本システムでは、カラーセット、テキスト、そしてウェブ検索時のランキング順位を判断基準として用いる(詳細は後述)。これは、カラリゼーションのアルゴリズムが、ソース画像内の形状、位置情報に左右されないからである。色空間には、3章において前述した Lab 色空間を用いる。これにより、色彩での画像検索を行いやすくし、さらに画像取得フェーズの段階で画像を Lab 色空間へ変換しておくことでカラリゼーションフェーズ時の合成時間を短縮させる。

本フェーズでは、最初に画像データベース内の全画像に対して Lab 色空間への変換をおこなう。変換後の値は、輝度チャンネル(L)とカラーセット(a, b)としてデータベースに格納する。さらに、カラリゼーションのソース画像としての有効性を測る基準として、各画像とソース画像とのテキスト類似度を測定し、

データベースに格納する。次に、データベース内において、類似したカラーセットごとにグループ化をおこなう（以下、類似カラーセットグループとする）。カラーセットの類似は、 $a, b$  を軸にした値の分散の類似をもとに判断する。そして、それぞれの類似カラーセットグループ内で、ウェブ検索でのランキング順位値が最も高いものをグループ代表画像とし、インタフェースにおいてプレビュー画像として表示させる。ユーザは、プレビュー画像の中から最も自分の主観に沿っていると感じる画像を選択する。選択された画像のカラーセットが、同章1節で定義している主観カラーセットとなる。画像が選択されると、今度はカラーセットグループ内で、テキスト類似度とランキング順位を重み付けした順位での並べ替えがおこなわれる。そして、上位から順に画像を用いて実際にカラリゼーションをおこない、ユーザに合成結果を提示していく。これにより、主観を保持したまま、ソース画像としてふさわしい画像を選択することができる（図4）。

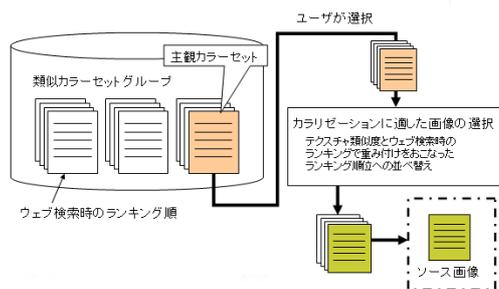


図4. ソース画像選択までの流れ

#### 4.4.対話型インタフェース

ユーザの主観を出力結果により強く反映させるため、対話性を重視したインタフェースを考える。これにより、ユーザは処理途中で試行錯誤を繰り返しながら理想の結果を求めることができる。本システムでは、以下の部分において、ユーザとの対話性をもたせてフィードバックを行う。各ポイントでは、自分の位置より前のポイントへ戻って作業しなおすこともできるようにする。

1. ウェブ画像検索の検索キーワードの入力
2. 主観カラーセットの選択
3. 最終的なソース画像の選択

#### 5.予備実験

検索エンジンを用いてウェブ画像検索を行う場合、検索結果は全て検索時に用いたキーワードに依存する。そのため、目的とする画像取得に向けて、どのようなキーワードを選択すべきであるかを検討しておく必要がある。本稿では、予備実験として Google イメージ検索を用いて、様々な検索キーワードでの検索結果の有効性の検証と考察を行った。

実験では、サンプルとして図5の3種類の画像を用いる。そして“それらの画像をカラリゼーション結果として出力させるような画像データベースを作るための画像”を検索するためのキーワードに関して具体的に検証を行う。検索されるべき画像の条件は、以下のように定義する。

- サンプル画像と同じ対象物が存在
- 色のもつ雰囲気が似ている
- 画面構成の類似は問わない

これを元に、様々なキーワードに対して、検索結果の上位100件中条件を満たす画像の数を調べる。検索結果が100件に満たなかった場合は、総検索数内での調査とする。また、検索画像の見つけやすさの指数として総検索結果数も調べる。



- (a) (b) (c)
- (a) 山：色領域がはっきりした単純な画像  
 (b) 猫：対象物がはっきりとした画像  
 (c) 町：複数オブジェクトにより構成された複雑な画像

図 5. 実験サンプル画像

● 単一キーワード

画像から連想されやすいキーワードを主要キーワードとよび、これらを単一で入力して検索結果について考察する。ここでは、主要キーワードを (a) 山・富士山・景色, (b) 猫・アメリカンショートヘア, (c) 町並み・ミハス・白い街, とする。

(a)	条件を満たす画像数(枚/100)	総検索結果数(件)
山	15	732,000
富士山	28	77,100
景色	6	96,900

(b)	条件を満たす画像数(枚/100)	総検索結果数(件)
猫	6	303,000
アメリカンショートヘア	36	1,200

(c)	条件を満たす画像数(枚/100)	総検索結果数(件)
町並み	2	30,600
ミハス	37	741
白い街	9	274

以上の結果から、キーワードとしては、より具体的な固有名詞を選択することが効果的であることがわかった。しかし、画像に対する主観が固まっていない場合、これらの具体的な名称を選択することは難しい。そこで、上位的な言葉をキーワードとし、検索結果からの選択の幅を広げた上で、じっくりと主観に沿う画像を探していくことも有効的であると考えられる。

これ以降は、主要キーワードを (a) 富士山, (b) アメリカンショートヘア, (c) ミハス, に絞って論ずる。

● 論理演算子 OR

Google 画像検索では、論理演算子を用いて検索キーワードを表現することができる。これを利用し、主要キーワードとその関連語を OR でつなげて、より

広い領域での検索を試みる。それぞれの関連語は、(a) 富士, Fuji (b) American Shorthair (c) アンダルシア, Mijas とする。

(a)	条件を満たす画像数(枚/100)	総検索結果数(件)
富士山	28	77,110
富士山 OR 富士	15	156,000
富士山 OR Fuji	10	500,000
富士山 OR 富士 OR Fuji	12	455,000

(b)	条件を満たす画像数(枚/100)	総検索結果数(件)
アメリカンショートヘア	39	1,200
アメリカンショートヘア OR "American Shorthair"	36	2,940

(c)	条件を満たす画像数(枚/100)	総検索結果数(件)
ミハス	37	741
ミハス OR アンダルシア	17	4,210
ミハス OR Mijas	17	13,400
ミハス OR アンダルシア OR Mijas	15	16,200

以上の結果より、OR を用いて複数のキーワードを指定した場合、主要キーワードを単一で入力した場合に比べて、総検索結果数は増加するが検索精度は低下してしまうことがわかった。これは、Google 画像検索のランキングアルゴリズムが要因になっていると推測される。たとえば、その他の条件がほぼ等しい状況下で「A OR B」の検索を行い「A」のみを含む画像と「A」と「B」の両方を含んだ画像が検索されたとき、後者のほうがランキング上位にくる、などの理由が考えられる。検索精度が低下することは問題ではあるが、総検索結果数が増加することで、より多くの可能性のある画像がピックアップされた点は非常に評価できる。そこで、本システムでは Google 画像検索の論理演算子 OR は用いない代わりに、関連語ごとに単一キーワード検索を行い、それぞれのランキング上位画像をまとめることによって、複数キーワードを用いた検索結果とする。

● バリエーション表現

ユーザが幅広いバリエーションでのカラリゼーションを行えるように、特別な主観をもつ状況下におけるキーワードの選択方法についても考察する。ここでは、(a) に対して、同じ対象物の夕焼けに染まったシーン (a') のカラリゼーションを想定する。そして、2 枚の画像に対するキーワードの比較をおこなう。



(a)



(a')

同一の主要キーワードを用いて，単一キーワード検索を行った結果，以下のようになった．

富士山	条件を満たす画像数(枚/100)
(a)	28
(a')	4

この場合 (a') の条件を満たす画像はほとんど見つからないことがわかる．そこで，(a) と (a') を比較した場合の差分を表すキーワード，もしくは (a') の状況を率直に表すキーワードの利用を考える．

- (a) と (a') の差分を表すキーワードとして「夕焼け」「朝焼け」などが予測される．これらキーワードのそれぞれを主要キーワードと AND でつないだ検索を行う．キーワードの決定には，文化的側面や科学的側面など様々な方向からの差分を考えることができる．これにより，ユーザもつ意識の方向性を考慮に入れた画像取得を実現できると考える．

富士山	条件を満たす画像数(枚/100)	総検索結果数(件)
AND 夕焼け	34	758
AND 朝焼け	18	462

- (a') の状況を表すキーワード「夕富士」などのキーワードが予測される．これを新しい主要キーワードとして，単一キーワード検索を行う．

(a')	条件を満たす画像数(枚/100)	総検索結果数(件)
富士山	4	77,110
夕富士	14	89

これらのことから，バリエーションをもつカラリゼーションを行う際には，より率直なキーワードを選択する，もしくは主要キーワードに差分となるようなキーワードをつなげた検索を行うことが有

効であると考えられる．

## 6.まとめと今後の課題

本研究では，ソース画像を用いて自動で着色を行うカラリゼーション手法に画像検索によるウェブ上からの画像取得を組み合わせた，新しいカラリゼーションシステムを提案した．これにより，最低限の操作で，ユーザの主観に沿った自然な色味でのカラリゼーションを行う．また，本稿ではこのシステムのための予備実験として，Google イメージ検索を用いてウェブ画像検索における検索キーワードの選択に関する考察も行った．

今後は，予備実験から得られた考察を元に，ウェブ画像検索フェーズのインタフェースの考案とシステムの実装を行っていく．画像類似検索フェーズにおいても，より技術的な観点からシステムを掘り下げ実装を目指す．

## 文 献

- 堀内 隆彦, "画素の種まきによるカラリゼーション," 画像電子学会論文誌, vol.33, no.1, pp.2-8, Jan.2004.
- Anat Levin, et al, "Colorization using Optimization," In Proc. SIGGRAPH2004, August 2004.
- Tomihisa Welsh, et al, "Transferring Color to Grayscale Images," Proc. SIGGRAPH2002, August 2002.
- E.Reinhard, et al., "Color Transfer between Images," IEEE Computer Graphics and Applications, pp.34--40, September/October 2001.
- <http://www.google.co.jp/>
- Tara Calishain, GOOGLE HACKS, 山名 早人 (監訳), pp.325-326, (株)オライリージャパン, 東京, 2003.
- 北川高嗣, 中西崇文, 清木康, "静止画像メディアデータを対象としたメタデータ自動抽出方式の実現とその意味的画像検索への適用," 情報処理学会論文誌: データベース, VOL.43, No.SIG12(TOD16), pp38-51, 2002.
- <http://mmm.nttrd.co.jp/>