

時間属性を考慮した視線の移動パターンの抽出

Extraction of Temporal Patterns from Eye-tracking data

江川 陽^{1*} 白山 晋^{1,2}
Akira Egawa¹, Susumu Shirayama^{1,2}

¹ 東京大学大学院工学系研究科, School of Engineering, University of Tokyo

² 東京大学人工物工学研究センター, RACE, University of Tokyo

Abstract: 人間の視線データを計測・分析する視線分析は、人間の高度認知処理の解明や、インターフェースの評価、知識の獲得など様々な分野において利用されている。しかしながら、視線データには、大量のノイズや個人差を含み、依然として視線の移動パターンの機械的な抽出などが難しい。そこで本稿では、視線データから時間属性を考慮した移動パターンの抽出を試みる。はじめに、視線計測の対象となる画像をその生来的な性質を利用して領域分割する。次に、計測された視線データを各領域上にマッピングすることにより、視線データを符号化する。そして、符号化された視線データに対して、ネットワーク分析手法の利用、被験者間の視線の類似度計算などを行い、移動パターンを分析する。被験者実験を通して提案手法の有効性を検証する。

Eye movement analysis is a popular method to analyze human cognitive process, evaluate usability of interface and extract implicit human knowledge. However, several issues exist in the conventional methods make quantitative analysis difficult. In this paper, we attempt to extract eye movement patterns in terms of temporal factors. First, eye movement data is acquired, and mapped into the image segmented by a saliency map model. Second, a network is generated by regarding the segments as nodes, and a quantitative measure of importance for the parts of the image is calculated by the PageRank algorithm. Third, temporal patterns of eye movement is extracted by calculating similarity between viewers' eye movement data. Usefulness of our method is discussed through an eye tracking experiment.

1 はじめに

人間の視線データを計測・分析する視線分析は、人間の高度認知処理の解明や、インターフェースの評価、知識の獲得など様々な分野において利用されている。例えば、Web ページ閲覧時の視線の動きを計測することにより、Web ページのどの領域が見られやすいか抽出する研究 [7]、あるいは、自動車運転時の着眼スキルの抽出と伝達を支援する研究 [11]、などがある。

しかしながら、一般的に視線追跡装置により測定される生の視線データは大量のノイズを含み、さらに個人差があるため、直接的に有用な情報を取得することは難しいとされる [2]。したがって、計測された視線データに対して何らかの後処理が必要となる。

視線データの後処理として代表的なものに、視線計測の対象となる視覚情報をいくつかの領域に分割し、その領域上に停留点を形成する手法がある [10, 11, 12]。この手法により、注目領域の抽出と視野の移動を分析するこ

とが可能となる。また、被験者の視覚情報への注目度を定量的指標として計算する手法がある。定量的指標の代表は、領域上で停留している注視点の個数（注視時間）にもとづく領域の注目度である。単純に加算する方法 [10]、パルツェン窓関数を用いるもの [11]、ガウス分布を重ねる方法 [5] などがある。

これらの手法は、いくつかの問題がある。前者には、領域分割に視線データの分析結果が大きく依存してしまう問題が挙げられる [8]。後者には、注目度の定量的指標計算において、領域の注視順序（視線が領域間をどのように移動したか）という情報が棄却されてしまっている。

我々のグループでは、領域分割の難しさと注目度の定量的指標の計算という二つの問題を解決するために、以下のような方法を提案している [9]。まずは、視線計測の対象となる画像の生得的な特徴に基づき領域分割し、その領域上に視線データをマッピングする。そして、領域間の視線の移動パターンをある種のネットワークと考えることで、各領域への注目度を算出するという手法である。しかしながら、この手法においても領域の注視順番のような時間属性が十分に考慮されていない。このた

*東京大学大学院工学系研究科
〒 277-8568 千葉県柏市柏の葉 5-1-5 総合研究棟 560 号室
E-mail: egawa@race.u-tokyo.ac.jp

め、被験者間において時系列情報を考慮した視線の移動パターンの比較, などの分析ができない。

そこで、本稿では、視線データの時間属性を考慮した視線の移動パターンの抽出を試みる。時間属性の中でも特に領域の注視順序に着目し、被験者間の注視行動の類似度を計算する手法を示す。そして、従来我々が提案している視線データ分析手法と比較し、その特性や有効性を検証する。

2 手法

2.1 入力画像の領域分割

2.1.1 Saliency map の作成

はじめに、Ittiらが提案しているSaliency mapモデル[3]にもとづき、視覚情報が有する人間が注意を集めやすい領域をsaliency mapとして抽出する。図1にsaliency mapの例を示す。図1左図が入力画像、右図がこの画像から作成されたsaliency mapである。Saliency mapは、入力画像の各領域がどれくらい注意を集めやすいかを定量化した画像であり、入力画像から一意に作成できる(ただし、いくつかのパラメータは存在する)。ここで、saliency mapを $s(l, m)$ とし、注意の集めやすさをsaliencyと呼ぶことにする。また、saliencyは0から255までの値をとるものとする。図1右図で白くなっている点はsaliencyが高いことを示している。

2.1.2 Saliency map のクラスタリング

Saliency map上で、saliencyが高い領域を抽出することでクラスタリングする。クラスタリングは、以下の2つのステップからなる。

はじめに、saliency map上で閾値よりもsaliencyが高い点を抽出し、表1にしたがいグルーピングする。擬似コードを表1に示す。 K はクラスタのインデックスを表す。



図1: Source image (left) and its saliency map (right). Photo by (c)Tomo.Yun [6]

また、クラスタ K の重心を $(x_K, y_K) = \frac{1}{N_K} (\sum_{(l,m) \in C_K} l \sum_{(l,m) \in C_K} m)$ とする。Saliency map 上の点 (l, m) が属するクラスタのインデックスを $c(l, m)$ 、 (l, m) の8近傍に存在するクラスタの中で (l, m) とクラスタ重心の距離が最も近いクラスタのインデックスを返す関数を $\text{NearestCluster}(l, m)$ とする。 δ は閾値であり、本稿では saliency map 上の最大 saliency の半分としている。このステップで抽出されるクラスタを、コアクラスタと呼ぶことにする。次に、コアクラスタに含まれる点の8近傍でコアクラスタに属していない点を、最も近いコアクラスタに割り当て、コアクラスタを拡大する。すべての点がいずれかのクラスタに属するまで同様の処理を行う。

このようにクラスタリングした saliency map を Saliency Cluster Map(SCM)と呼ぶ。SCMの例を、図2に示す。図2左図はSCMにおいてコアクラスタのみを示しており、右図はコアクラスタ以外の領域も示している。

2.1.3 SCM を利用した画像の分割

SCMを利用して画像のセグメンテーションを行う。SCMにおけるクラスタと対応する画像上の領域をセグメントとして抽出する。この際、SCMのコアクラスタに対応する領域をコアセグメントとして扱う。図1のセグメンテーション結果の例を、図3の左図に示す。セグメントの中でコアセグメントに対応する領域を濃い色、その他の領域を薄い色で示している。

表1: Pseudocode for constructing the *core-clusters*

```

set 0 to  $c(l, m)$ 
 $K \leftarrow 0$ 
for  $l = 0$  to  $[(imax - 1)/16]$ :
  for  $m = 0$  to  $[(jmax - 1)/16]$ :
    if  $s(l, m) \geq \delta$  and  $c(l, m) = 0$  {
      if at least one of the eight neighbors of  $(l, m)$ 
        belongs to any cluster {
         $K \leftarrow K + 1$ 
         $c(l, m) = K$ 
         $(x_K, y_K) = (l, m)$ 
      }
    }
    else {
       $k = \text{NearestCluster}(l, m)$ 
       $c(l, m) = k$ 
      update  $(x_k, y_k)$ 
    }
  }
  if  $s(l, m) < \delta$  {  $c(l, m) = 0$  }

```

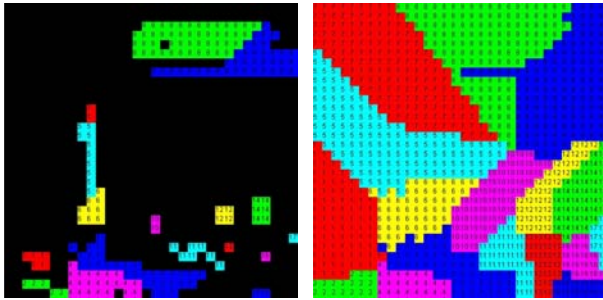


図 2: Saliency Cluster Map.

2.2 初期重要度マップの作成

Itti らは、視覚的注意 (FOA: Focus Of Attention) の移動モデルも提案している [3]。このモデルでは、FOA は saliency の高い領域から低い領域へ順番に移動する。本研究では、Itti らと同様に、この移動を視覚情報の持つ生来的な性質に起因する視線の移動と考え、各コアセグメントにおける saliency の最大値を利用して、FOA の移動を計算し、移動の順番に応じてセグメントの順位付けを行う。k 番目の FOA として選ばれたセグメントのうち、コアセグメントを R_k 、コアセグメント以外の領域を R'_k という指標で表す ($k = 1, \dots, M$)。図 3 左図においては、1 番目と 2 番目の FOA として計算されたセグメントを示している。

次に、FOA の移動を利用して重要度を計算する。セグメント R_k の重要度を i_k^1 とし、 i_k^1 を以下の式で求める。

$$i_k^1 = \frac{1}{k \cdot \sum_{k=1}^M \frac{1}{k}}. \quad (1)$$

本稿では、FOA の移動と上式で求めた各セグメントの重要度を付与した図を初期重要度マップと呼ぶ。図 3 右図に初期重要度マップの例を示す。初期重要度マップは、人間の視知覚モデルである Saliency map モデルをもとに、領域の注目度と注目のされ方を定量化したものである。

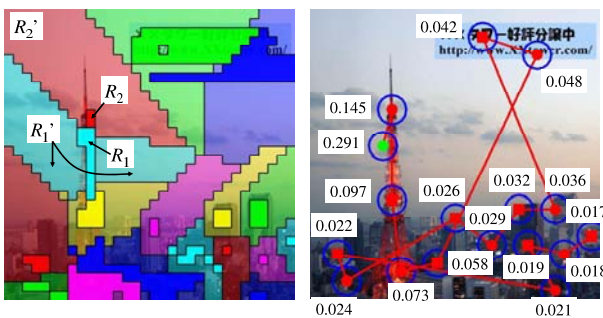
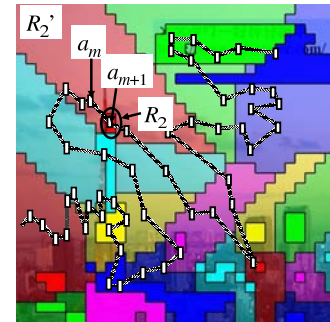


図 3: Example of image segmentation(left) and first Image Importance Map(right).



$A^j = (...a_m, a_{m+1}...) \rightarrow (...R_2', R_2...)$

図 4: Schema of encoding the eye-movement dataset.

2.3 視線データの符号化

初期重要度マップは、決定論的なモデルである Saliency map モデルをもとに作成されるので、個人固有の見方は反映されていない。そこで、注視行動の個人差が反映される定量的指標を検討する。

はじめに、視線計測装置を利用して、視覚情報に対する被験者の視線データを計測する。n 番目の視線データを、 $a_n = (x_n, y_n, t_n)$ と表す。 (x_n, y_n) は a_n の座標であり、 t_n は計測された時間を表す。個人 j の視線データの系列を、 $A^j = (a_1, a_2, \dots, a_N)$ と表す。N は視線データの総数とする。第 2 に、2.1 節で述べたセグメント上に A^j をマッピングし、視線データをセグメントの指標によって符号化する。図 4 に示すように、 a_m と a_{m+1} がそれぞれ R_2' 、 R_2 上に位置するとき、これらの視線データは $R_2'R_2$ として符号化される。

この符号化された視線データは、どの領域をどのような順番で注視したかという情報を表す。このデータに各領域の注視時間を加えたものをガントチャート風に可視化した図が図 5 である。縦軸は各セグメントを表し、横軸は時間を表している。この図からは、人間の視線行動には、見比べる、一つの領域を長時間注視する、順番ずつ注視するなどの何らかのパターンが存在していることが示唆される。よって、このようなチャートを利用する

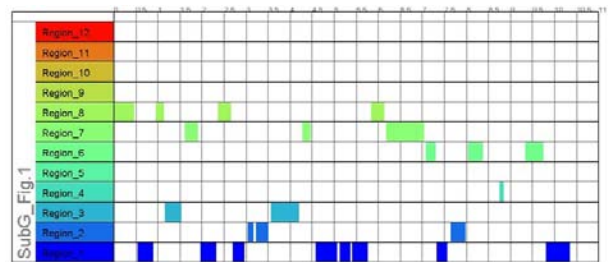


図 5: Visualization of eye movement pattern.

ことで、被験者の注視行動の詳細な分析が可能となる。しかし、ある被験者独自の視線の移動パターンの抽出、あるいは被験者間の移動パターンの比較などをする際には、その被験者の視線の移動パターンの特徴を示す何らかの定量的な指標が必要となる。そこで、ここからは視線データの定量的な指標を計算するために、いくつかの手法を示す。

2.4 第2の重要度マップの作成

本節では、視線の移動データをネットワーク（有向グラフ）と考え、ネットワークの位相構造をもとに、各セグメントの定量的な重要度を算出する。

まずは、コアセグメント R_k の重心をノード v_k 、セグメント間の視線の動きをリンク（アーク）とすることで、符号化された視線データからネットワークを形成する。本稿では、この際に、コアセグメント外のセグメントに位置するものをコアセグメントに置き換えた。この点は今後の課題である。

次に、ネットワークの各ノード（セグメント）の重要度を PageRank アルゴリズムを利用して算出する。はじめに、視線のネットワークの隣接行列を行ベクトルによる推移確率行列に変換する。

例えば、 $R_2R_4R_1R_1R_1R_2R_4R_3$ と符号化された視線データからは、 $v_2 \rightarrow v_4 \rightarrow v_1 \rightarrow v_1 \rightarrow v_1 \rightarrow v_2 \rightarrow v_4 \rightarrow v_3$ というネットワークが形成される。リンクの数を重みとするループも含む隣接行列 A からは、以下の推移確率行列 T が得られる。

$$T = \begin{matrix} & v_1 & v_2 & v_3 & v_4 \\ \begin{matrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{matrix} & \begin{pmatrix} 2/3 & 1/3 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \end{pmatrix} \end{matrix} \quad (2)$$

この推移確率行列を利用して、各ノードの重要度を計算する。ノード v_k の PageRank を i_k^2 、全体を $\mathbf{p} = (i_1^2, \dots, i_k^2, \dots, i_M^2)$ とすると、視線ネットワークの各ノードの PageRank は以下のように求められる。

$$\mathbf{p}^{v+1} = T^v \mathbf{p}^v \quad (3)$$

ここで、 v は反復回数を示す。このように符号化された視線データからネットワークを作成し、ノードに相当するセグメントに PageRank を付与したものを第2の重要度マップと呼ぶ。第2の重要度マップの例を図6に示す。図中の矢印は視線データから作成したネットワーク、数字は PageRank の値であり、これが各セグメントの重要度になる。第2の重要度マップは Saliency map モデルという視知覚モデルにもとづき分割された領域の重要性を、注視された領域の移動という観点から定量化したものである。

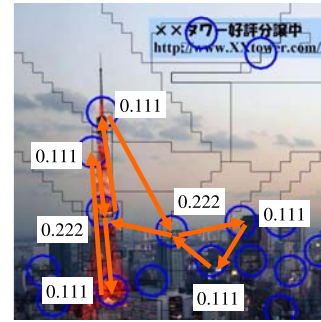


図 6: Second Image Importance Map.

2.5 重要度マップの統合

初期重要度マップと第2の重要度マップを統合し、領域毎の重要度を算出する。これを統合重要度と呼び、セグメント R_k の統合重要度を i_k^s とし、全体を $\mathbf{p} = (i_1^s, \dots, i_k^s, \dots, i_M^s)$ で表す。統合重要度を、biased PageRank アルゴリズム [4] にもとづき、次式で求める。

$$\mathbf{p}^{v+1} = \omega T^v \mathbf{p}^v + (1 - \omega)\mathbf{i} \quad (4)$$

ここで、 $\omega (0 \leq \omega \leq 1)$ は統合パラメータである。本稿では、 $\omega = 0.5$ としている。

2.6 視線の移動パターンの類似度の算出

以上までに示してきた一連の重要度マップ作成手法は、視線の移動パターンにもとづいた各領域の重要度を定量的に示すものである。しかし、これらの定量的指標においては、各領域の注視順序や注視時間といった視線データが本来有している時系列情報を考慮していない（図5を参照）。特に注視順序は、領域を見比べる、ある一定の順序に沿って領域を注視する、などのパターンを考える上で重要であると考えられる。そこで、本節では、被験者間における領域の注視順序を顕在化させるために、符号化された視線データ間の類似度を計算する。

符号化された視線データは、被験者がどのような順番で対象画像を見たかという情報を含む文字列である。このような視線データに対して Brandt らは、任意の二つの文字列間の編集コストを利用して、視線データ間の類似度を計算している [1]。本稿においても、Brandt らの手法を利用して、被験者間の符号化された視線データの類似度を計算する。

ここでは、二つの文字列間の編集コストを、Levenstein 距離を利用して算出する。Levenstein 距離とは、ある二つの文字列について、文字の挿入や削除、置換の操作によって、一つの文字列を別の文字列に変形するのに必要な手順の最小回数として与えられる。

二人の被験者 j と j' の符号化された視線データ G_j と $G_{j'}$ の Levenstein 距離を $LD(G_j, G_{j'})$ とすると, Brandt らによれば二人の視線データの類似度 $Simirality(G_j, G_{j'})$ は,

$$Simirality(G_j, G_{j'}) = 1 - \frac{LD(G_j, G_{j'})}{Stringlength(G_j, G_{j'})} \quad (5)$$

と表現できる. ただし, $Stringlength(G_j, G_{j'})$ は文字列 G_j と $G_{j'}$ の長さのうち長い方の値である. 例えば, 二つの視線データ $R_1R_1R_2R_4R_3$ と $R_2R_1R_5R_3R_4R_3$ の類似度は 0.5 である.

この類似度を求めることにより, 視線の時系列での移動順序を考慮した比較が可能となる.

3 実験

本章ではこれまでに述べた視線データの分析手法の有効性を検証するために, ある画像に対して視線計測実験を行う. 対象となる画像は, 図 7(a) 左図で示す流体现象のシミュレーション結果の可視化画像である. この画像に対して, 3 名の被験者の視線データを計測した. 被験者のうち一名は流体力学の専門家(被験者 A)で, 残り 2 名は非専門家の大学生(被験者 B, C)である. 被験者には画像を自由にってもらい, データ分析には画像提示から 10 秒間の視線データを利用した.

3.1 重要度マップ

図 7(a) 右図に対象画像の初期重要度マップ, 同図 (b), (c) に被験者ごとの第 2 の重要度マップ, 統合重要度マップをそれぞれ示す.

流体力学の専門家である被験者 A によれば, 図 7(a) 左図において赤い丸で囲まれている二つの領域がこの流体现象において重要であるという. 被験者 A はこれらの領域を, 一方の領域の初期重要度が低いにもかかわらず注視している. よって, 二つの領域の第 2 の重要度ならびに統合重要度が高く算出されている.

被験者 B は, これら領域のうち片方を少し注視したに過ぎず, この領域の重要度が低く算出されている. 被験者 C も二つの重要領域のうち片方のみを注視したが, 被験者 B と比較して頻繁に注視して. これは, 注視領域の初期重要度 (saliency) が高く, 被験者の注意を集めやすかったからであると考えられる.

このように一連の重要度マップを作成することによって, 各領域の定量的な重要度をもとにした視線データの分析が可能となる.

表 2: 視線データの類似度

	被験者 A	被験者 B	被験者 C
被験者 A	1	0.185	0.286
被験者 B	0.185	1	0.321
被験者 C	0.286	0.321	1

3.2 視線の移動データの類似度

三人の被験者の視線データの類似度を表にまとめたものが表 2 である. これより, 被験者 A と被験者 B, C 間の類似度に注目する. 被験者 A と B の類似度が他の二者間の類似度に比べて低く出ている. これは, 両者の注視行動が大きく異なっていたことを示している. 前節で示したように, 被験者 A は対象の左側の重要領域を中心に, 被験者 B は対象の中心部を中心に注視するという傾向の違いが類似度に反映したと思われる. また, 被験者 A と被験者 C の類似度が被験者 A と B の類似度と比べて高く出ているのは, 二つの重要領域のうちの片方を両者とも頻繁に注視していたからであると考えられる.

このように, 注視順序を考慮した被験者間の視線データの類似度を利用することで, 被験者間の視線の移動パターンの違いを定量的に分析することが可能となる. 一方, 視線データの類似度指標のみでは, 被験者 A と C が重要領域を互いに注視していた, といった詳細な分析は困難である. よって, 視線データの類似度と重要度マップを利用した領域の重要度を単独で利用するのではなく, これらを組み合わせることによって, より詳細な注視順序を考慮した視線データの分析が可能となると考えられる.

4 まとめ

本稿では, 視線データの時間属性を考慮した視線データからの移動パターンの抽出をするために, 時間属性の中でも特に領域の注視順序に着目し, 被験者間の注視行動の類似度を計算する手法を示した. 加えて, 被験者実験を通して我々が従来から提案している重要度マップを利用した視線データ分析手法と比較し, その特性や有効性を検証した. 結果, 二つの分析手法を組み合わせることで, 詳細な視線データの分析が可能となることを示した.

今後は本稿で取り扱わなかった領域の注視時間を考慮した視線データの分析手法などを行いたい.

参考文献

- [1] S. A. Brandt and L. W. Stark. Spontaneous eye movements during visual imagery reflects the content of

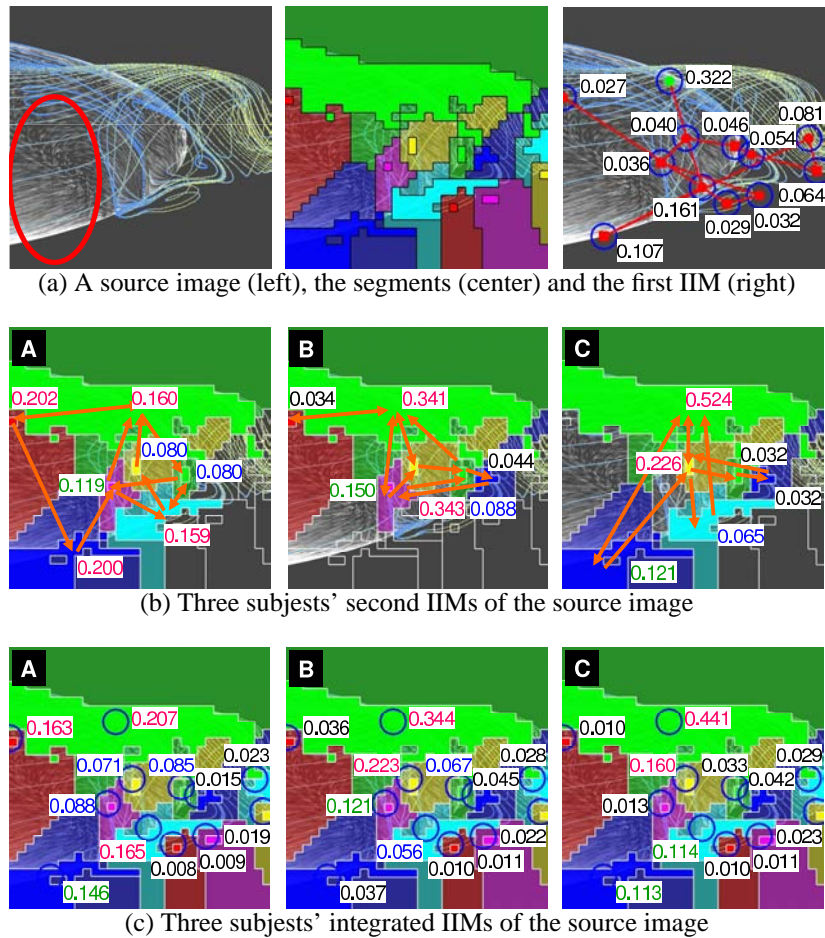


図 7: Results for the experiment.

- the visual scene. *Journal of Cognitive Neuroscience*, Vol. 9, No. 1, pp. 27–38, 1997.
- [2] A.T. Duchowski. *Eye Tracking Methodology*. Springer, second edition, 2007.
- [3] L. Itti, C. Koch, and E. Niebur. A Model of Saliency-Based Visual Attention for Rapid Scene Analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20, No. 11, pp. 1254–1259, 1998.
- [4] S.D. Kamvar, T.H. Haveliwala, C.D. Manning, and G.H. Golub. Exploiting the Block Structure of the Web for Computing Pagerank. *Stanford University Technical Report 2003-17*, 2003.
- [5] O. Špakov and D. Miniotas. Visualization of Eye Gaze Data using Heat Maps. *Electronics and Electrical Engineering*, Vol. 74, No. 2, pp. 55–58, 2007.
- [6] ゆんフリー写真素材. <http://www.yunphoto.net/>.
- [7] 高久雅生, 寺井仁, 江草由佳, 齋藤ひとみ, 三輪眞木子, 神門典子. Web 情報探索における視線データの予備的分析. *情報知識学会誌*, Vol. 18, No. 2, pp. 181–188, 2008.
- [8] 江川陽, 白山晋. 注目点からの効率的・効果的な画像コンテンツの再構成手法. *映像情報メディア学会誌*, Vol. 62, No. 9, pp. 1443–1452, 2008.
- [9] 江川陽, 白山晋. 視線分析手法の高度化とその応用. 第 23 回人工知能学会全国大会論文集, 2009.
- [10] 大野健彦. 視線から何がわかるか - 視線測定に基づく高次認知処理の解明. *認知科学*, Vol. 9, No. 4, pp. 565–576, 2002.
- [11] 藤本武司, 砂山渡, 山口智浩, 谷内田正彦. 視線行動の可視化による着眼スキル伝達支援. *人工知能学会論文誌*, Vol. 19, No. 3B, pp. 174–183, 2004.
- [12] 脇山孝貴, 吉高淳夫, 平嶋宗. 注目の検出に基づいた興味モデルの作成と絵画推薦. *情報処理学会論文誌*, Vol. 48, No. 3, pp. 1048–1057, 2007.