G-06

階層化 NFTG と夜間歩行者姿勢推定への応用

Nighttime Pedestrian's Posture Estimation using Hierarchy NFTG

前渕 啓材†　　呉 海元　　和田 俊和
Hiroki Maebuchi　　Haiyuan Wu　　Toshikazu Wada

Abstract

本稿では，近赤外線カメラで撮影された人物画像から夜間歩行者の姿勢推定を目的とする．事前に用意された複数の姿勢を持つ人物輪郭モデルと入力画像のエッジ画像との類似度を測定し，最も類似度の高いモデルを人物画像の姿勢とする．しかし，全モデルと比較を行うと処理時間が掛かる．この問題を解決するため，本稿では，NFTG(Nearest First Traversing Graph)[2]を用いて人物輪郭モデルの相互関係をグラフ化する．また，姿勢推定におけるグラフ探索の場合，「初期値依存性」と「モデル数をどのくらいにするか」の 2 つの問題点が挙げられる．これらの問題点に対処するため，本稿では，階層化 NFTG を提案する．

1．はじめに

夜間歩行者と自動車との交通事故の死亡率は，昼間と比べ，5 倍高いという調査結果が得られている．それに伴い，近赤外線カメラや遠赤外線カメラ等を使って，夜間における歩行者を検知し，夜間走行のドライバーの視覚を支援できる歩行者検出システムの開発が進められている[4]-[10].

また，車と歩行者の接触事故において，道沿いに歩く人物より横断する歩行者の方が交通事故が多い．そこで，本稿では，事故発生の危険度である歩行者姿勢方向を推定する．



図 1.1 歩行者姿勢推定

近赤外線画像は人物の輪郭が検出しやすい特性がある．提案手法では，事前に様々な歩行姿勢における多数の人物輪郭モデルを用意し，姿勢が未知である入力画像に対して，エッジ検出処理を行い，得られたエッジ画像を使い，人物輪郭モデルと類似性を測定する．最も類似度の高い人物輪郭モデルの姿勢を入力画像の姿勢とする．入力画像と人物輪郭モデル間の類似性を測るためには，距離尺度が必要である．エッジ画像から類似性を求める方法として，Hausdorff distance や Chamfer distance がある．

続いて，姿勢を探索する効率性について考える．夜間走行のドライバーの視覚支援システムとして利用することを想定しているため，夜間歩行者の姿勢推定は精度と速度を両立する必要がある．精度を保証するため，様々な歩行姿勢の人物輪郭モデルをできるだけ多く用意する必要がある．一方，多数の人物輪郭モデルに対して，すべて入力画像との距離尺度を測っていては，効率が悪い．そこで，本稿では，人物輪郭モデルをグラフ化し，必要最低限の探索を行う．しかし，グラフ探索の場合，モデルの枚数をいくつに設定するか，といった問題がある．少なければ，当然精度は悪く，多く用意すれば，処理時間が掛かるといったトレードオフの問題が起きる．また，初期値問題も発生する．そこで，本稿では，グラフ構造を階層化し，グラフの中から幾つかの代表点を選択することで，上階層ではノード間の距離が大きく，大雑把な探索を行った後，下階層で，より正確な探索を行うため，多数のノードが存在する構造を提案する．

本論文の構成は以下の通りである．2 章で Hausdorff distance や Chamfer distance の距離尺度について述べる．3 章では，グラフ構造の作成方法である NFTG について述べ，4 章では，その NFTG を階層化する方法について述べる．5 章では，それぞれの実験結果について記述し，6 章でまとめる．

2．エッジ画像を用いた距離測定

エッジ画像同士の距離を測定する方法として，Hausdorff distance と Chamfer distance がよく利用される．それぞれの距離変換画像を作成し，エッジからの最短距離を測る．2.1 章で Hausdorff distance，2.2 章で Chamfer distance について説明する．エッジ画像，距離変換画像，エッジ画像間の距離測定方法について，図 2.1 に示す．

2.1 Hausdorff distance

Hausdorff distance は任意のエッジ画像 A と画像 B から図形の形状間の相違度を測定する手法である．本論文では，画像 A を入力画像のエッジ画像とし，画像 B を歩行者の人物輪郭モデルとする．

画像 A，B から抽出したエッジ情報を $A=\{a_1,a_2,\cdots,a_p\}$ $B=\{b_1,b_2,\cdots,b_q\}$ とする．この時，集合 A と集合 B の Hausdorff distance は以下の(2.1)式，(2.2)式，(2.3)式で定義することができる．

†和歌山大学,
Faculty of Systems Engineering Wakayama University

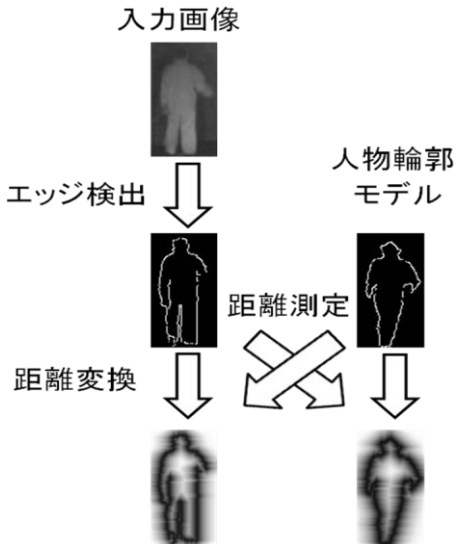


図 2.1　エッジ画像間の距離測定

$$h(A,B)=\max_{a\in A}\min_{b\in B}\|a-b\| \tag{2.1}$$

$$h(B,A)=\max_{b\in B}\min_{a\in A}\|b-a\| \tag{2.2}$$

$$H(A,B)=\max(h(A,B),h(B,A)) \tag{2.3}$$

$h(A,B)$を集合 A から集合 B までの最大最小距離（フォーワード・ディスタンス）, $h(B,A)$ を集合 B から集合 A までの最大最小距離（リバース・ディスタンス）とする．フォーワード・ディスタンス，リバース・ディスタンスの 2 方向間の距離からの最大値 *H(A,B)*を Hausdorff distance とする．

Hausdorff distance は 2 方向から測定することによって，形状間が大幅に異なるものでも形状間が若干しか変化していないものでも相違度を画素単位で検出することが可能である．

しかし，Hausdorff distance はエッジ点からエッジ点までの画素間の距離しかない．そのため，正解付近になると相違度の分解能が低くなるという問題点もある．

2.2　Chamfer distance

Chamfer distance は(2.4)式，(2.5)式で定義され，エッジ画像間の相違度を求めることができる．

$$C(A,B)=\sum\min_{b\in B}\|a-b\| \tag{2.4}$$

$$C(B,A)=\sum\min_{a\in A}\|b-a\| \tag{2.5}$$

入力画像のエッジ画像と人物輪郭モデルをそれぞれ重ね合わせた時，入力画像のエッジ点から人物輪郭モデルの最も近いエッジ点までの距離の総和を計算する．Hausdorff distance 同様，距離変換を用いることにより，簡単に計算することができる．

Hausdorff distance と比べ，Chamfer distance は集合間の距離の総和を測定する方法であるため，相違度の分解能が高い．

2.3　グラフ構築用距離尺度

Hausdorff distance と Chamfer distance の比較結果を表 5.1 に示す．推定時間，推定精度ともに Chamfer distance の方が良いことを確認した．これらの結果に基づいて，本提案手法の距離尺度には Chamfer distance を採用する．

3．近さ優先探索グラフ(Nearest First Traversing Graph)

2 章では，人物画像の姿勢推定する方法について述べた．3 章からは出来る限り高速に姿勢推定する方法を提案する．ある人物画像に対して，全ての人物輪郭モデルと比較を行う必要は無い．類似する人物輪郭モデル同士を辺で繋ぎ，人物画像との距離が近い方へと辿るグラフ構造を作ることで，探索回数は削減できる．本章では近さ優先探索グラフ(Nearest First Traversing Graph：NFTG)について述べる．

3.1　NFTG の構成方法

NFTGは，「クエリはグラフの頂点中にある」と仮定すると，「任意の頂点をクエリとしたNstFT(Nearest First Traversing)の手続きが，頂点以外の点で停止しない」という特徴がある．この特徴を満たすためには以下の性質が必要となる．

連結性：全ての頂点は連結している．
単調性：任意の頂点からNstFTで頂点vcを移動させるたびに，クエリqとvcとの距離は単調に減少する．
極小性：グラフのどの一辺でも取り除くと上記の二つの性質が満足されない．すなわち，冗長性のないグラフでなければならない．

これらの性質のうち，単調性が満足されれば，連結性は満足される．なぜならば，連結でなければ単調性は満足されないからである．また，単調性を満たさないグラフにある辺を追加することによって単調性（と連結性）が満足されたとすると，少なくとも冗長なグラフにはなっていないと言える．グラフ構築時に追加する辺がある頂点に関する単調性を成立させるために必要不可欠な辺であれば，このようにして出来上がるグラフは極小性も満足する．したがって，任意の頂点に関して単調性を満足させることが本質であり，順に辺を追加していき，この性質が成立した時点で辺の挿入をやめれば，上記の 3

つの性質を満足するグラフが得られる．

単調性を満足するために必要となる補題を以下に示す．

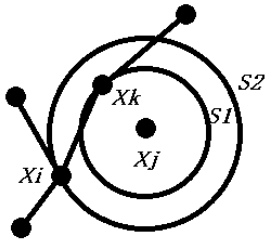


図3.1　NonStuckとStuck

[補題１]
頂点Xiを起点としXjをクエリとするNstFTがXiで停止しない条件は，Xjを中心とし，Xiを通過する超球の内部にXiから一つの辺で繋がった頂点が存在することである．

[補題２]
Xjを中心としXiを通過する超球の内部にXiから１つの辺で繋がった頂点が無い場合，頂点Xiを起点としXjをクエリとするNstFTはXiで停止する．

２つの補題から次の定理が成り立つ．

[定理１]
あるグラフG=(V,E)が次の条件を満足する時，Gは単調性を満足する．

$$\forall v_i, v_j \in V, NonStuck(v_i, v_j) \qquad (4.1)$$

この定理から，NstFTに用いることが出来るグラフが満たすべき条件が明らかになった．実際には，Stuck(Vi,Vj)という性質を満たす頂点間に辺を挿入することを繰り返すことによって定理１の条件を満たせばよい．

3.2　Directed NFTG

距離の公理を以下に示す．

1. 同一性　$d(p,q)=0 \quad \Leftrightarrow \quad p=q$

2. 対称性　$d(p,q)=d(q,p)$

3. 三角不等式　$d(p,r)\le d(p,q)+d(q,r)$

4.非負性　$d(p,q)\le 0$

Chamfer distance は d(p,q)≠d(q,p)となり，対称性を満たさない．したがって，距離の公理を満たしていないことになる．これは，エッジ画像と距離変換画像間で距離測定を行ったためであると考えられる．

NFTG は無向グラフで構成されているため，Chamfer distance で NFTG を構成するためには有向グラフ(Directecd NFTG : DNFTG)を作成しなければならない．そこで，構成時には，1 つの頂点に対して，エッジ画像と距離変換画像の 2 種類の画像を登録し，「Chamfer distance はエッジ画像から距離変換画像への距離」という定義に基づいて方向に注意しながら，画像間の距離比較を行えばよい．

4．NFTG の階層化について

近赤外線画像から歩行者姿勢を推定するため，人物輪郭モデルを使って，グラフ探索を行う場合，初期値と人物輪郭モデルの枚数は探索精度と処理時間に影響する．初期値をランダムに与えた時，目的頂点と離れていた場合，距離測定を多数行わなければならない．また，NFTG を構成する上で，モデルを何枚用意するか，といった問題も存在する．例えば，モデルを必要以上に用意した場合，それだけグラフは大規模になり，距離測定回数が多くなる．しかし，モデルが少なければ，適切な姿勢推定ができない．

本論文では，初期値とモデル枚数の問題を解決するため，階層化 NFTG グラフを提案する．具体的に，NFTG を階層化し，下階層から少数代表的な候補点を選択して，上階層の少ない人物輪郭モデルでグラフを作成し粗い探索を行う．下階層では，その結果を初期値として与え，より細かく正確なグラフ探索を行う．階層化 NFTG の構成図について図 4.1 に示す．黒丸は各ノードとし，円柱をそれぞれの階層とする．

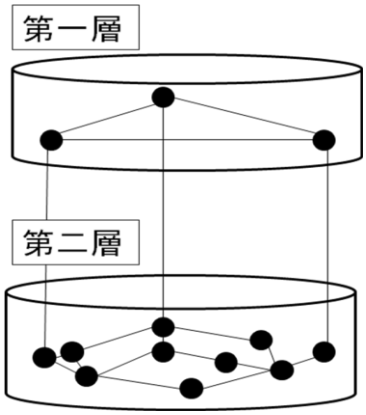


図 4.1 階層化 NFTG

4.1　階層化 NFTG における上階層の候補点選択

下階層からの代表となる候補点の選択方法について述べる．候補点は頂点間同士がある程度以上離れていなければ，意味が無い．そこで，最大距離クラスタリングを用いて，候補点 **P** を選択する．

$$\mathrm{P} = \arg\max_{a\neq b\in(1\cdots n)} \min\{d(A,B)\} \qquad (4.1)$$

Chamfer distance は距離の公理を満たしていない．したがって，頂点 A から頂点 B までの距離は離れているが，頂点 B から頂点 A までの距離が非常に近いといった事態が発生する．そのような事態を避けるため，候補点を選ぶ際，(4.2)式のように 2 つの距離の内，近くにある距離を頂点間の距離とする．

$$D(A,B)=\min\{d(A,B),d(B,A)\} \qquad (4.2)$$

5． 実験

本章では，Hausdorff distance と Chamfer distance との精度比較，姿勢推定の結果，探索効率についての実験結果を記載する．

5.1 Hausdorff distance と Chamfer distance の比較実験

1 章，2.2 章で述べた Hausdorff distance と Chamfer distance の姿勢推定における精度，処理時間について表 5.1 に示す．表 5.1 から処理時間，正解率ともに Chamfer distance の方が良いことを確認した．

表 5.1 各距離における処理時間と精度の比較結果

距離	処理時間	正解率
Hausdorff	29.7μsec	47%
Chamfer	11.7μsec	62%

続いて，Hausdorff distance と Chamfer distance によって，最も相違度が高く，入力人物の姿勢と似ていると判断された人物輪郭モデルを図 5.1 に示す．

姿勢毎の推定結果の精度比較結果を図 5.2 に示す．横軸は姿勢の種類（歩行者が前後，右，左に歩いている画像，歩行者が自転車に乗り，前後，右，左を走行している画像），縦軸は正解率である．図 5.2 から各姿勢の正解率は Chamfer distance の方が高いことを確認した．

5.2 2 次元における階層化 NFTG の実験結果

階層化 NFTG の有効性について確認するため，まず，2 次元データで階層化 NFTG を構築した．構築された階層化 NFTG を用いて平均探索回数と分散値を求めた．NFTG を構成する上での頂点数は 10000 個とし，頂点の値は 0 から 10000 までとする．また，距離尺度はユークリッド距離とする．


図 5.1(a)


図 5.1(b)


図 5.1(c)


図 5.1(d)


図 5.1(e)


図 5.1(f)

図 5.1(a)(d) 入力画像，図 5.1(b)(e)Hausdorff distance によって，選択されたモデル 図 5.1(c)(f)Chamfer distance によって，選択されたモデル

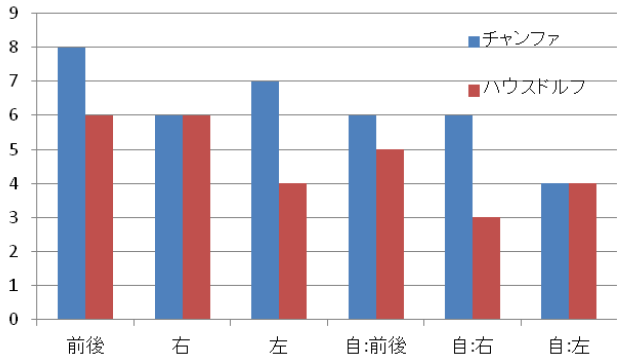


図 5.2 Chamfer distance と Hausdorff distance との推定精度の比較結果

実験結果を図 5.3(a)と図 5.3(b)にそれぞれ示す．縦軸がそれぞれ，平均探索回数，分散値，横軸が両方とも上階層の頂点数である．階層化しない場合(横軸が 0)と比べ，平均探索回数が 3 倍ほど少なくなっている．また，分散値も大幅に落ち，安定した探索が可能となっている．

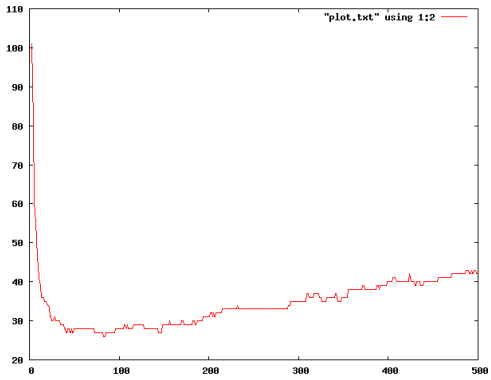


図 5.3(a)階層化 NFTG を用いた平均探索回数

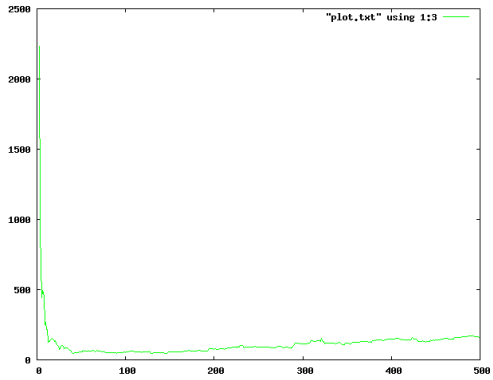


図 5.3(b)階層化 NFTG を用いた分散値

5.3　人物輪郭モデルにおける階層化 NFTG の実験結果

続いて，提案手法を用いて，実際の人物輪郭モデルで階層化 NFTG を構築した．

構築された階層化 NFTG を使って，平均探索回数と分散値を求めた．実験結果について図 5.4(a)，図 5.4(b)にそれぞれ示す．縦軸がそれぞれ，平均探索回数，分散値，横軸が上階層の頂点数である．

人物輪郭モデルを使った場合，探索において平均探索回数，分散値共に 2 次元での平均探索回数，分散値よりも探索効率が落ちている．これは扱うデータが高次元になったことが原因と考えられる．しかし，平均探索回数，分散値ともに候補点が 400 付近で最も低くなっており，安定かつ高速の探索が出来ていることが分かる．

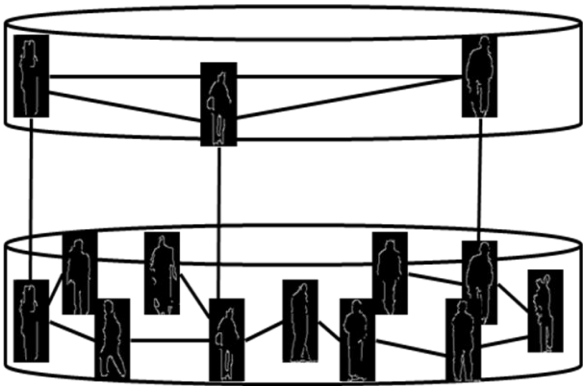


図 5.4 人物輪郭モデルを用いた階層化 NFTG

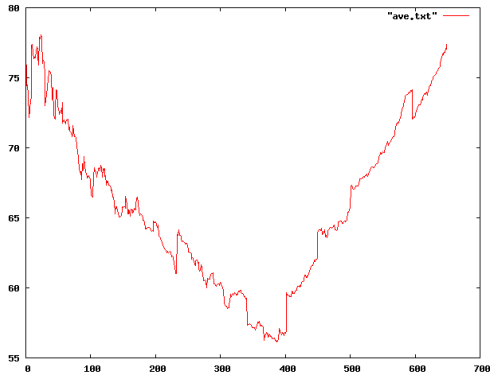


図 5.4(a)階層化 NFTG を歩行者姿勢推定における平均探索回数

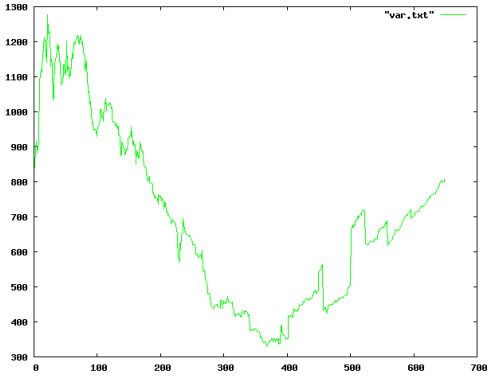


図 5.4(b)階層化 NFTG を歩行者姿勢推定における分散値

6．　おわりに

本稿では，夜間における歩行者の姿勢推定と探索方法について述べ，Chamfer distance を用いた階層化 NFTG を提案した．近赤外線画像を用いた実験結果より，グラフ構造である NFTG を階層化することによって，安定かつ高速に探索できることを確認した．

謝辞

本研究の一部は，文部科学省科学研究費補助金基盤研究(c) 24500205 の補助を受けている．

参考文献

[1] 前渕啓材，呉海元，和田俊和，幾何制約付 HOG と NFTG による夜間歩行者検出,電子情報通信学会技術研究報告, 信学技報, PRMU2011-175, pp.291-296, 2012

[2]Junya S , Toshikazu W , “Nearest First Traversing Graph for Simultaneous Object Tracking and Recognition” , IEEE Computer Vision and Pattern Recognition , pp.1-7 , 2007

[3]Sergey Brin, “Near Neighbor Search in Large Metric Spaces” , VLDB pp.574-584 , 1995
[4] Broggi,A, Fedriga, R.L, Tagliati,A, “Pedestrian Detection on a Moving Vehicle: an Investigation about Near Infra-Red Images, ” Intelligent Vehicles Symposium, pp431-436 Tokyo Japan, Jun.2006
[5] Hui Cao, Koichiro Yamaguchi, Takashi Naito, and Yoshiki Ninomiya, “ Pedestrian Recognition Using Second-Order HOG Feature, ” ACCV, pp.439-449 Queenstown NewZealand Nov.2010
[6] Thi Thi Zin, Pyke Tin, and Hiromitsu Hama, “Bundling Multislit-HOG Features of Near Infrared Images for Pedestrian Detection, ” Innovative Computing, Information and Control (ICICIC) , pp302-305, Kaohsiung Taiwan, Dec.2009
[7] Junfeng Ge, Yupin Luo, Gyomei Tei, “Real-Time Pedestrian Detection and Tracking at Nighttime for Driver-Assistance Systems, ” Intelligent Transportation Systems, IEEE Transactions on, vol.10, no.2, pp.283-298, June.2009
[8] Jianfei Dong, Junfeng Ge, Yupin Luo, “Nighttime Pedestrian Detection with Near Infrared using Cascaded Classifiers,” Image Processing, 2007. ICIP2007.IEEE International Conference on, vol.6, pp.VI-185-VI-188, Oct.2007
[9] Ronan O’Malley, Edward Jones, Martin Glavin,” Detection of pedestrians in far-infrared automotive night vision using region-growing and clothing distortion compensation”, Infrared Physics & Technology, vol.53, pp.439-449, Nov.2010
[10] Soga M, Hiratsuka S, Fukamachi H, Ninomiya Y, “Pedestrian Detection for a Near Infrared Imaging System” , ITSC2008. 11th International IEEE Conference on, pp.1167-1172, Oct.2008
[11]David Vignon, Brian Lovell,”Real-Time Hausdorff-Based Tracking”,Digital Image Computing Techniques and Applications Jan.2002
[12]William J,”Efficiently Locating Objects Using the Hausdorff Distance”,International Journal of Computer Vision 24(3) pp.251-270,1997
[13]Yonggag Jin,Fazin Mokhtarian,Miroslaw Bober,John Illingworth,”Fuzzy Chamfer Distance and its Probabilistic Formulation for Visual Tracking”,CVPR,pp.1-8,Aug.2008
[14]Tianyang Ma, Xingwei Yang, Longin Jan Latecki, “Boosting Chamfer Matching by Learning Chamfer Distance Normalization” , ECCV,pp.450-463, 2010
[15]曽暉, 大西昇, 田中敏光, 杉江昇”近似 Chamfer 距離とカーネル部分空間を用いたシルエット画像からの三次元人物姿勢推定”, 電子情報通信学会研究報告 pp.53-58, Feb, 2007
[16]Ramanan Navaratnam, Andrew W.Fitzgibbon, Roberto Cipolla, “The Joint Manifold Model for Semi-superised Multi-valued Regression”, ICCV, pp.1-8, 2007
[17]大西克則, 滝口哲也, 有木康雄, “HOG 特徴に基づく単眼画像からの人体 3 次元姿勢推定”, 画像の認識・理解シンポジウム(MIRU2008), IS3-16, pp.960-965, 2008
[18]Bingbing Ni, Gang Wang, Pierre Moulin, “RGBD-HuDaAct: A Color-Depth Video Database For Human Daily Activity Recognition”, ICCV, pp.1147-1153, 2011
[19]S. Arya, D. Mount, N. Netanyahu, R. Silverman, and A. Wu. “An optimal algorithm for approximate nearest neighbor Searching”, Journal of the ACM, pp.891–923, 1998
[20]D.Barnea, H.Silverman. “A class of algorithms for fast digital image registration”, IEEE Trans. Computers, 21 pp.179–186, 1972
[21]J. Bentley, “Multidimensional binary search trees used for associative searching”, Comm. of ACM, 18 pp.509–517, 1975.
[22]M. J. Black and A.Jepson, “Eigentracking: Robust matching and tracking of articulated objects using a view-based representation”, IJCV, 26, pp.63–84, 1998
[23]T. Bozkaya and M. Ozsoyoglu. Distance-based indexing for high dimensional metric spaces. 1997 ACM SIGMOD, 1997