

# 道路周辺の3次元点群データを用いた 車載カメラ画像からの車両の位置姿勢推定

黒部 聡亮<sup>1,a)</sup> 中島 雅貴<sup>1,b)</sup> 原 孝介<sup>2,c)</sup> 木下 久史<sup>3,d)</sup> 齋藤 英雄<sup>1,e)</sup>

**概要:** 近年、高齢者による車やバスによる交通事故が多発しており多くの犠牲者が出ている。しかし、現在それらの事故現場の検証は、ブレーキ痕や周辺の破損物の状況から予想をして、事故当時の復元を行うのが主な方法である。しかし、この方法には、事故以前の道路の状態等に影響を受けてしまい正確に予測ができないだけでなく、定量的に事故時の状況を3次元復元することはできないという問題点が存在する。本研究では、道路周辺の3次元情報をレーザースキャナにより色情報付3次元点群として保存しておき、それを用いて任意のカメラの画像からその3次元位置姿勢の復元を推定する手法を提案する。本手法は、軌跡の復元を行いたい、連続したドライブレコーダ画像を入力とし、初期値を設定することで自動的に各フレームにおける車両の位置姿勢の推定を行う。推定は、3次元点群から生成した候補画像とドライブレコーダ画像とのマッチングを利用し、各フレームの推定結果を次フレームの粗推定値として利用することで、連続したフレームに対して処理を行う。実験結果では、本手法によりカメラの位置姿勢の軌跡を高い精度で推定することができたことを示す。

**キーワード:** 車両位置姿勢推定, 自動車, 3次元点群

## 1. はじめに

近年、高齢者による車の交通事故が多発しており多くの犠牲者が出ている。しかし、現在それらの事故現場の検証は、ブレーキ痕や周辺の破損物の状況から予想をして、事故当時の復元を行うのが主な方法である。しかし、この方法には、事故以前の道路の状態等に影響を受けてしまい正確に予測ができないだけでなく、定量的に事故時の状況を3次元復元することはできないという問題点が存在する。

また、交通事故の発生を抑制する目的の運転支援システム [1][2] や事故回避の自動ブレーキ [3][4]、そして自動運転 [5] に関する研究が盛んに行われている。これらの研究は車両が事故を起こすことを事前に防ぐために構築されたシステムである。それらのシステムを実現させるに際し、事故を起こす直前直後の車両の描く軌跡は貴重なデータの1つであり、それらをコンピューターに学習させることで同様な軌跡を描いた車両に自動ブレーキをかけるといった

制御をかけることで事故を防ぐ、様々な応用が期待される。

そこで、本研究では将来の実用化を見据え、事故が多発している交差点や路地を高精度なレーザースキャナを用いて3次元点群として取得しておき、ドライブレコーダ画像を入力した際に、そのドライブレコーダ画像が撮影された瞬間の車両の位置姿勢を推定する手法を提案する。本手法では、取得した3次元点群から候補画像を複数枚生成し、それらとドライブレコーダ画像の比較を行い、マッチング数と距離の側面から最も適切な候補画像を決定し、最終的な車両の位置姿勢推定を行うという手順を経ることで、正確な結果を得ることを可能としている。先行研究としては Yungpeng ら手法 [6] や Radu らの手法 [7] がある。これらの手法は大規模な3次元点群内に対して、ある画像がどの位置姿勢で撮影されたかを推定するもので、3次元点群内のランドマークの認識や3次元特徴量による手法を用いている。本研究は、実際に事故を起こした車両に搭載されているドライブレコーダの画像が撮影された位置姿勢を高精度に推定することが必要なため、毎フレーム点群から画像を生成し、ドライブレコーダ画像とマッチングしている。

評価実験では、2つの実験を行なった。1つ目は、1枚の入力画像に対して車両の位置姿勢を推定する実験、2つ目は連続した25フレームを入力として、各フレームにおける車両の位置姿勢を推定し、描いた軌跡を復元する実験で

<sup>1</sup> 慶應義塾大学

<sup>2</sup> 株式会社デンソーアイティラボラトリ

<sup>3</sup> 株式会社デンソー

a) kurobe@hvrl.ics.keio.ac.jp

b) nakashima@hvrl.ics.keio.ac.jp

c) khara@d-itlab.co.jp

d) HISASHI\_KINOSHITA@denso.co.jp

e) hs@keio.jp

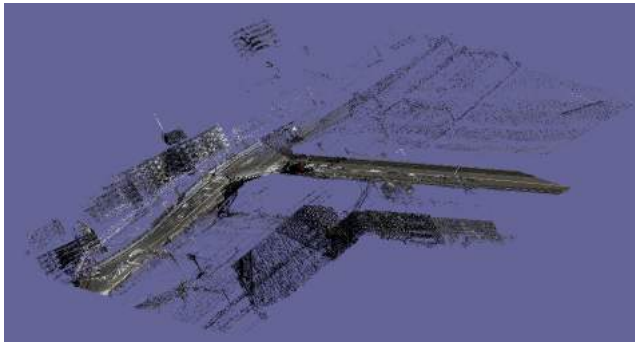


図 1: 取得した 3 次元点群

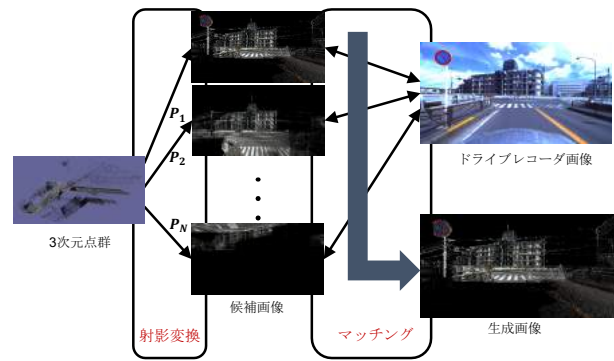


図 2: 生成画像決定の概念図

ある。2つ目の実験においては、1 フレーム目のみに初期値を与え、現在のフレームの位置姿勢推定結果を次フレームの粗推定値として利用する連続的な処理における、本手法の有効性の確認を行った。

## 2. 提案手法

本提案手法は、4つのパートにより構成される。1つ目は、事故が多発する道路周辺の情報をレーザスキャナを用いて3次元点群として取得するパート、2つ目は、取得した3次元点群に射影変換を施して候補画像を生成するパート、3つ目は、生成した複数の候補画像の中から実際の位置姿勢推定に使用する画像の選択するパート、4つ目は、生成した画像とドライブレコーダ画像のマッチング結果から車両の位置姿勢を推定するパートである。本章ではそれらのパートを詳細に説明する。

### 2.1 3次元点群の取得

まず本手法では、道路周辺の情報をレーザスキャナを用いて、色付3次元点群として取得する。取得した点群のスケールは現実の絶対スケールと同じ比率であり、この点群を今後の処理で利用する。

### 2.2 候補画像群の生成

本研究の目的は、実際に事故が起きた際に現場周辺の点群情報とドライブレコーダ画像から車両の描いた軌跡を復元することにある。位置姿勢の推定を行いたいドライブレコーダの1シーンを決定し、その画像が撮影された瞬間の車両の位置情報をGPSから取得する。しかし、このGPSの情報には数メートル単位の誤差が乗っている。そこで、本手法では、このGPSの位置情報を元に複数の位置姿勢の候補を作り出し、その位置姿勢に相当する射影変換を点群に施すことで候補画像を複数生成する。点群を画像平面に投影する処理では、同じピクセルに複数の点が投影される可能性が考えられる。そこで、本手法では、車両に最も近い3次元点の投影を最終的な生成画像に反映させるため、車両を原点とした座標系において、進行方向の距離情

報を考慮して投影を行う処理を行なっている。この処理によって、実際にはドライブレコーダ画像に写っていない物体や影などが生成画像に含まれることを軽減することができる。

### 2.3 位置姿勢推定に用いる画像の選択

2.2節の処理によって、GPSの情報から候補画像を複数枚生成されている。これらの画像全てと位置姿勢の推定を行いたいドライブレコーダ画像を総当たりでマッチングを行う。マッチングの際に、対応点と対応点の距離を全て算出する。このマッチングを行う際にロバスト推定法のRANSAC[8]を用いた。

生成した複数枚の候補画像とドライブレコーダ画像とのマッチングには、AKAZE特徴量[9]を使用した。AKAZE特徴量はKAZE特徴量[10]の処理速度を向上させたアルゴリズムを使用している。SIFT特徴量[11]やSURF特徴量[12]で使用されているGaussian filterによるスケールベースはGaussian filterが周囲の画素に等方的であるのでオブジェクトのエッジをぼかしてしまい、局所的な特徴を抽出できないという欠点がある。KAZE特徴量は、その欠点を解決するために、非線形で非等方的なスケールベースを使用することで、局所的な特徴も抽出することができるように処理が組まれている。しかし、KAZE特徴量は計算に時間がかかるという欠点がある。そこでAKAZE特徴量では、Feature Descriptorとして、Modified-Local Difference Binary (M-LDB)という独自のDescriptorを使用し、さらにピラミッド構造の計算を高速化するための独自の工夫を組み入れることで、ロバスト性の向上と高速化を図った特徴量である。

本手法では、AKAZE特徴量により抽出した特徴点に対して、flannのアルゴリズム[13]を用いてマッチングの探索を行なっている。その後RANSACを用いて誤対応の削減を行っている。また、マッチングの距離の閾値を大きくするとマッチング数は増えるが誤対応が増え、閾値を小さくするとマッチング数は減るが誤対応が減るというトレードオフの関係がある。

後の処理で方程式を解く関係から、対応点が最低6点以上の条件で候補画像を絞る。そして、残った候補画像の中で最も対応点の距離の平均が小さい候補画像を最終的な推定に利用する。処理の流れを図2に示した。

## 2.4 車両の位置姿勢推定

2.3節の処理によって、GPSの情報を元に点群から生成した複数枚の生成画像の中から、最終的な位置姿勢の推定に利用する画像を選択した。生成された画像は点群に射影変換を施し、画像平面に投影することで生成しているのでマッチングに成功した対応点の3次元位置は、射影変換の逆演算を行うことで計算することが可能である。選択された生成画像と位置姿勢推定の行いたいドライブレコーダのAKAZEによるマッチングにより、画像座標と点群における世界座標の対応を6点以上取得することができる。

$$P = A(R|t) \quad (1)$$

$$R = \begin{pmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \phi & \theta & \sin \phi \\ 0 & 1 & 0 \\ -\sin \phi & 0 & \cos \phi \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix} \quad (2)$$

$$t = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (3)$$

$$A = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \quad (4)$$

取得した対応点をRANSACによって誤対応を減らし、射影変換行列  $P$  を求める。 $P$  は式1に示すようにカメラの内部パラメータ  $A$  と車両の姿勢を表す  $R$ , 位置を表す  $t$  で構成されている。 $A$  はドライブレコーダのカメラキャリブレーションを行うことで求めることができるので、 $A^{-1}$  を式1の両辺の左から掛け、 $(R|t)$  を計算することで車両の位置姿勢を推定する。

## 3. 評価実験

本章では提案手法を評価するために行った実験及びその結果と考察について述べる。実験は2つ行なった。1つ目は、1枚のドライブレコーダ画像を入力として、その画像が撮影された瞬間の車両の位置姿勢を推定する実験。2つ目は、複数枚のドライブレコーダ画像を入力として、初期値を与え各フレームの推定値を次フレームの粗推定値として用いることで連続的に軌跡を推定する実験である。位置はRTKGPSとの比較、姿勢に関しては推定した位置姿勢から画像を生成し、推定したドライブレコーダ画像と目視による評価を行なった。

## 3.1 評価環境

本実験の実験環境は次の通りである。CPU: Intel Core i7-5820K 3.30GHz, GPU: GeForce GTX 960, RAM: 64GB.

1つ目の実験では、図3のように、それぞれ異なる2枚のドライブレコーダ画像の位置姿勢の推定を行なった。それぞれのドライブレコーダ画像のtimestampの情報から、車両に搭載されているGPSの位置情報を元に本手法を適応した。

2つ目の実験では、図4のように連続した25枚の画像を入力としている。毎フレームのGPS情報を取得し、角度を変化させて画像を生成する処理を行うと処理時間が膨大に掛かってしまうため、実用化に向けて1枚目に初期値を与えて本手法により位置姿勢の推定を行い、その結果を次フレームの初期値として利用する処理を25枚の画像に適応することで本手法の有効性を確認した。



図3: 実験1の画像 (左: シーン1 右: シーン2)



連続シーンの25枚

図4: 実験2の画像

## 3.2 評価結果

### 3.2.1 実験1

表1, 表2に各シーンの推定結果とRTKGPSから得られる真値との誤差を示した。位置に関しては、高い精度で位置情報を取得できるRTKGPSとの差から本手法の有効性の確認を行う。シーン1に関しては、X,Y,Zの全ての向きに置いて数センチメートル程度の誤差で推定できていることがわかる。シーン2は、X,Yに置いては10センチメートル程度の誤差で推定できているが、Z方向には50センチメートル程度の誤差が生じている。この原因として考えられるのは、Z軸が車の進行方向を表すことが考えられる。図3のシーン2の画像を見ると、進行方向に似たような風景が続いている。これは、世界座標と画像座標の関係を求める時に生じる誤差が、Z軸方向に大きくなってしまふことが考えられる。

姿勢に関しては、RTKGPSから真値を取得することが

表 1: シーン 1 の推定結果と真値との誤差

座標	測定結果 (m)	RTKGPS による真値 (m)	誤差 (m)
X	-366.25	-366.32	0.07
Y	13.45	13.49	0.04
Z	-440.73	-440.68	0.05

表 2: シーン 2 の推定結果と真値との誤差

座標	測定結果 (m)	RTKGPS による真値 (m)	誤差 (m)
X	-346.56	-346.57	0.01
Y	5.83	5.93	0.10
Z	-255.74	-255.22	0.52

できないため、精度の評価は推定した位置姿勢を元に、点群に射影変換を施して生成した画像とドライブレコーダ画像の比較を行い目視により推定結果を評価する。図 5, 図 6 に推定結果から生成した画像を示した。シーン 1, シーン 2 の生成画像を見るとドライブレコーダ画像が撮影された際の姿勢についても正確に推定できているのが分かる。以上から、シーン 1, シーン 2 の位置姿勢どちらにおいても正確に推定することができることを示した。



図 5: シーン 1 の推定結果  
(左: ドライブレコーダ画像, 右: 推定結果)



図 6: シーン 2 の推定結果  
(左: ドライブレコーダ画像, 右: 推定結果)

### 3.2.2 実験 2

実験 1 では 1 枚のドライブレコーダ画像を入力として、車両の位置姿勢を推定を行なったが、この方法を連続したシーケンスに適応すると処理時間が膨大に掛かってしまうので、実験 2 では 1 枚目のみに初期値を与えて、残りの画像には前フレームの推定結果を元にして処理を継続した結果を、毎フレーム RTKGPS と比較して精度の検証を行った。

図 7~図 9 に各フレームの推定値とそれに対応する RTKGPS の値を示した。どの座標においても正確に推定できていることが分かる。表 3 に各座標の誤差の平均を示した。

表 3 を見ると、X,Y,Z 座標の全てにおいて 1 メートル以

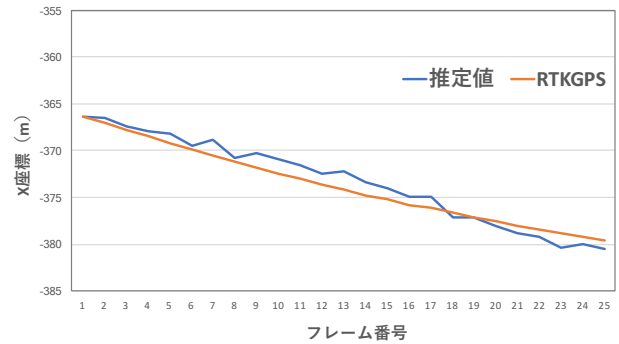


図 7: 実験 2 の推定結果 (X 座標)

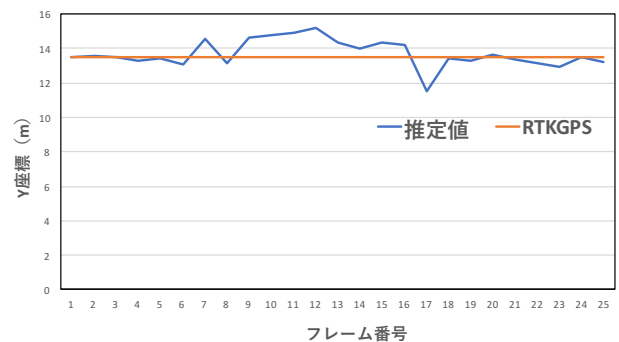


図 8: 実験 2 の推定結果 (Y 座標)

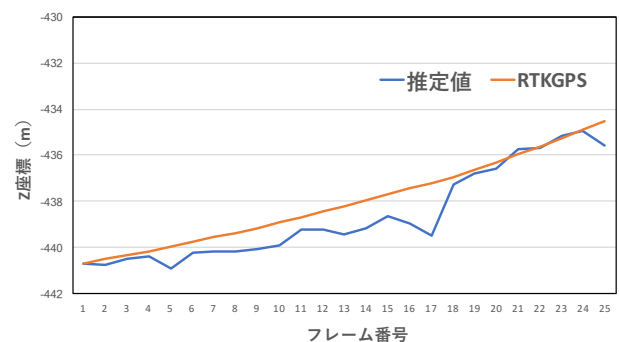


図 9: 実験 2 の推定結果 (Z 座標)

表 3: 25 フレームの誤差の平均

座標	誤差の平均 (m)
X	0.924
Y	0.579
Z	0.643

内の誤差で推定できていることが分かる。連続のシーケンスに対しても車両の位置を正確に推定できていることが分かる。姿勢に関しても、実験 1 と同様に、推定した位置姿勢を元に点群に射影変換を施し生成した画像とドライブレコーダ画像を比較することで、精度の評価を行う。図 10 に、推定した 25 枚のうち 8 枚の推定結果を示した。どの

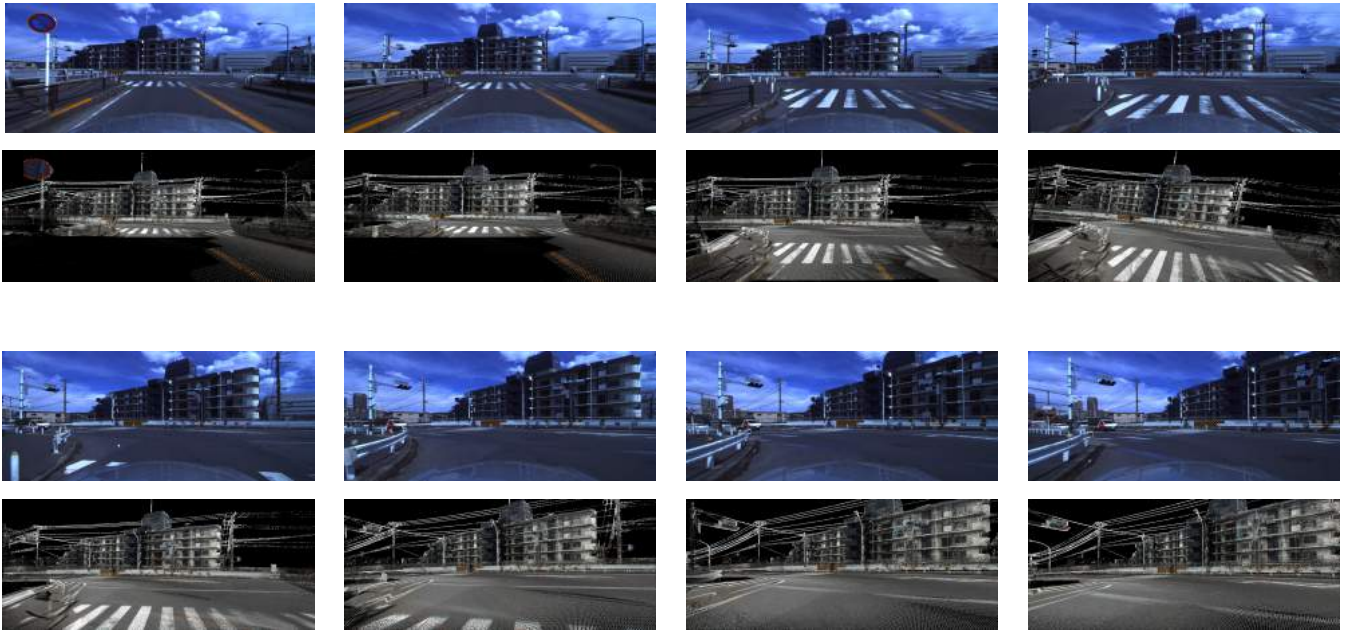


図 10: 推定結果 (1 段目, 3 段目: ドライブレコーダ画像, 2 段目, 4 段目: 対応する推定結果)

フレームにおいても, 概ね正確に推定できていることが分かる.

#### 4. 結論

本稿では, ドライブレコーダ画像を入力として, 取得した 3 次元点群から候補画像を生成し, その画像とドライブレコーダ画像のマッチングを通して車両の位置姿勢を推定する方法を提案した. 評価実験では, ドライブレコーダ画像を 1 枚を入力とするケースと, 連続した 25 枚の画像を入力するケースで本手法の有効性の確認した. 1 枚を入力とする実験では数 10 センチメートルの誤差で推定できること, 25 枚を入力とする実験では 1 フレーム目のみに初期値を与え, それ以降のフレームに関しては, 前フレームの推定値を利用することで連続的に正確な処理を行い, 軌跡を復元できること, 25 フレームを通し平均して 1 メートル以内の誤差で位置姿勢の推定ができるを確認した.

謝辞 本研究は, 経済産業省「平成 28 年度スマートモビリティシステム研究開発・実証事業 (事故データベースの構築技術の開発)」による成果である.

#### 参考文献

[1] M. Doi, K. Zeng, T. Wada, S. Doi, N. Tsuru, K. Iasaji and S. Morikawa: Steering-assist control system on curved road using car-to-car communication; International Conference on Intelligent Transportation Systems, pp. 1 - 6, 2016.

[2] G. David, A. M. Lopez, A. D. Sappa: Survey of pedestrian detection for advanced driver assistance systems; IEEE transactions on pattern analysis and machine intelligence, pp. 1239 - 1258, 2010.

[3] T. Bo, C. Stanley, H. Zhi and C. Yaobin: Pedestrian

protection using the integration of V2V and the Pedestrian Automatic Emergency Braking System; International Conference on Intelligent Transportation Systems, pp. 2213 - 2218, 2013

[4] C. Erik, A. Eidehall and Mattias Bengtsson: Collision warning with full auto brake and pedestrian detection-a practical example of automatic emergency braking; Intelligent Transportation Systems, pp. 155 - 160, 2010.

[5] R. Sasaki and S. Yasunobu: An intelligent auto-driving system by interactive acquisition of driving knowledge as information on route; International Conference on Intelligent Transportation Systems, pp. 179 - 182, 2004.

[6] Y. Li, N. Snavely, D. P. Huttenlocher and P. Fua: World-wide pose estimation using 3d point clouds; Springer International Publishing, pp. 147-163, 2016.

[7] R. B. Rusu, G. Bradski and R. Thibaux: Fast 3d recognition and pose using the viewpoint feature histogram; Intelligent Robots and Systems, pp. 2152 - 2162, 2010.

[8] M. A. Fischler and R. C. Bolles: Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography; Communications of the ACM, Vol. 24, No. 6, pp. 381 - 395, 1981.

[9] P. F. Alcantarilla, J. Nuevo and A. Bartoli: Fast Explicit Diffusion for Accelerated Features in Nonlinear Scale Spaces; British Machine Vision Conference, 2013.

[10] P. F. Alcantarilla, A. Bartoli and A. J. Davison: KAZE Features; European Conference on Computer Vision, pp. 214 - 227, 2012.

[11] D. G. Lowe: Distinctive image features from scale-invariant keypoints; International Journal of Computer Vision, Vol. 60, pp.91 - 110, 2004.

[12] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool: Speeded-Up Robust Features(SURF); Computer Vision and Image Understanding, pp. 346 - 359, 2008.

[13] M. Marius and David G. Lowe: Fast approximate nearest neighbors with automatic algorithm configuration; VIS-APP, pp. 331 - 340, 2009.