

# アドホック視覚センサネットワークにおける 視野接続推定

## Vision Graph Estimation for an Ad hoc Vision Sensor Network

近藤真, 鏡慎吾, 橋本浩一

Shin Kondo, Shingo Kagami, Koichi Hashimoto

東北大学大学院情報科学研究科システム情報科学専攻

Department of System Information Sciences, Graduate School of Information Sciences,  
Tohoku University

キーワード: アドホックネットワーク(ad-hoc network), 視野接続(vision graph, vision topology),  
SIFT(Scale Invariant Feature Transform),

連絡先: 〒980-8579 仙台市青葉区荒巻字青葉6-6-01 東北大学 情報科学研究科 橋本・鏡研究室  
近藤真, TEL 022-795-7021, FAX 022-795-7019, E-mail: kondo@ic.is.tohoku.ac.jp

### 1. 緒言

無作為に配置されたセンサが自律的にネットワークを構築するアドホックセンサネットワークの研究は,これから到来するユビキタス社会において重要なテーマである.

一方でカメラを用いたセンサネットワークの分野では,複数のカメラをネットワークで結び協調動作させる研究が盛んに行われている<sup>1), 2)</sup>.ただし,カメラネットワークは静的に構成されていることが多く,アドホックネットワークを想定していることは少ない.

またカメラの分野では近年集積回路技術の向上によりカメラの小型化,低コスト化,低電力化がますます進んでおり,従来では考えられなかった応用が実現可能になってきている,カプセル型の胃カメラなどの極小の無線カメ

ラの開発が進んでおり,極小カメラを大量にバラまくように配置させるアドホックな視覚センサネットワークの実現も現実味を帯びてきている.人の立ち入りが困難な災害現場でカメラを散布し負傷者を見つけるなど,視覚センサの新たな応用が期待されている.

このようなアドホックな視覚センサネットワークを構築するには従来のアドホックネットワークとは別の難しさがある.通常のアドホックセンサネットワークでは,図1の(a)のように物理的に隣接したセンサ同士が互いに接続されていく.しかし,視覚センサの場合は,センサ同士が物理的に隣接していても共通の事象を見ているとは限らない.2つのセンサが共通の事象を観測しているかどうかという図1の(b)のような視野接続関係を推定する必要がある.

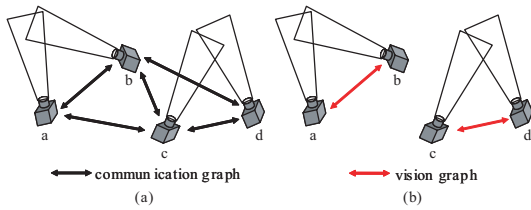


Fig. 1 視野接続関係と物理的接続関係

視野接続関係に着目した研究は従来から行われているが，その多くでは接続関係を手動で，あるいは何らかの便宜的な事前知識に基づいて行っていた<sup>3), 4)</sup>．一方いくつかの最近の研究では，画像内の動き情報に基づいて視野接続推定<sup>5)</sup>や，画像間の特徴点の対応づけに基づいて推定<sup>6)</sup>も行われている．本論文では，Cheng ら<sup>6)</sup>と同様に特徴点の対応づけに基づく視野接続推定を提案する．

## 2. 視野接続の推定

### 2.1 視野接続推定の原理

視野接続推定を行うためには2つの画像間で共通の事象を観測しているかどうかを調べる必要がある．本論文ではその推定方法として，Cheng ら<sup>6)</sup>同様にSIFT (Scale Invariant Feature Transform)を用いる．SIFT<sup>7)</sup>は画像中から特徴点を見つけ，その点を表現する特徴量を画像の回転やスケール倍にロバストな形で得ることができるものであり，画像認識等で広く用いられている．SIFTにより得られる特徴量は標準的には128次元のベクトルであり，特徴点の対応づけはその128次元空間中でのユークリッド距離が近い特徴点を探すことで行われる．ただし，SIFTでは誤った対応点も検出されるので，本論文では，画像間で一定数以上の対応点が検出されたら視野が接続していると判断する．

### 2.2 視野接続推定の分散化

多数のカメラノードの相互間の視野接続を推定する場合は，全カメラノード対全カメラノードについて特徴点の対応づけを行う必要がある．それを行う一つの方法として，すべてのカメラが持つデータを一つのノードに集め，そのノードですべての対応づけを計算する方法が考えられる．しかし，そのノードの周囲のリンクに負荷が集中してしまうことや，そのノードが機能しないと対応づけができないことなどの問題があり，一般にセンサネットワークには適さない．

また別の方法としては，各カメラノードのデータを他のカメラノードすべてに送信（ブロードキャスト）した上で，各カメラノードにおける視野接続先はそのカメラノード自身に計算させるという方法が考えられる．しかしこの方法では，同じデータをカメラノード数の数だけ重複して送るので無駄が生じ，ネットワーク全体のデータ通信の負荷が大きくなる問題がある．その解決策としてCheng ら<sup>6)</sup>は，データをブロードキャストするが，特徴点数の削減，特徴量次元の圧縮によりデータ通信量を削減することでネットワーク全体の負荷を低減している．しかし，重複による無駄が生じる点には変わりがない．

## 3. 提案手法

### 3.1 基本方針

既に述べたように，SIFT 特徴点の対応づけは，その特徴量間のユークリッド距離が近いものを探索することによって行われる．よって，対応づけられる特徴点对は，特徴量空間内で互いに近くに位置していると考えられる．この考察から，特徴量空間を適当に分割してお

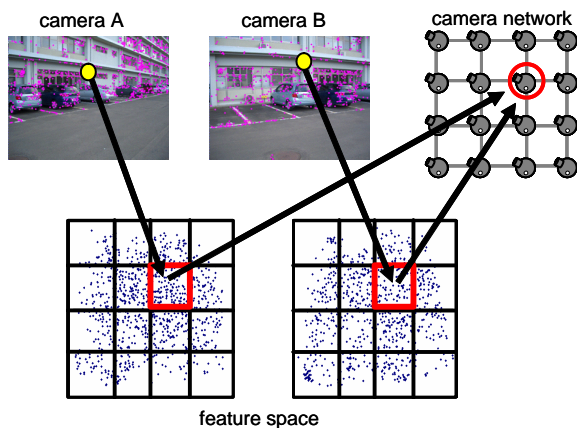


Fig. 2 提案手法の概要

き，探索処理はその分割された領域内のみに対して，その領域を担当するノードに実行させるという考えが導かれる．従来はあるノードで検出された特徴点情報を他のすべてのノードにブロードキャストしていたところを，その特徴量が含まれる分割領域を担当するノードに送るだけですむようになるため，ネットワークの負荷を低減できると期待できる．また，対応づけ計算の負荷分散も期待できる．この考え方を図 2 に示す．

ここで考えている対応点探索の分散化の考え方は，Peer-to-Peer (P2P) ネットワークにおける資源探索の分散化<sup>8)</sup>と類似している．例えば P2P ネットワークによるファイル共有アプリケーションでは，多数のネットワークノードの中から目的のファイルを保有するものを高速に見つけ出す手段が必要である．これを効率よく分散化して行う方法として分散ハッシュテーブルと呼ばれる技術が用いられる．分散ハッシュテーブルでは，ファイルの内容をハッシュ関数により数値化し，その数値とノードを対応づけることで，そのファイルの管理情報を持つノードを決定する．センサネットワークの分野では，センサデータの種類等を表すメタ情報をハッシュ化することでデータの保存先を決定するデータセントリックストレ

ジ<sup>9)</sup>等への応用が知られている．

これらと我々の問題の間の大きな違いは，我々は画像特徴量というセンサ情報を，ノード決定のためのキーとして直接用いようとする点にあり，従って完全一致による探索ではなく近傍探索が目的になる点にある．そのため，探索空間を複数の領域に分割することで，本来見つかるべき対応点が見つからなくなる恐れがある．

一般に，探索空間の分割を細かくすれば負荷分散の効果は上がるが探索ミスの可能性は大きくなると考えられる．ミスを低減するためには，ある程度は領域分割境界をまたいだ探索も行わなくてはならない．そのために必要なオーバヘッドをできるだけ減らすためには，特徴量空間内で近くに位置する分割領域どうしについては，それらの領域を担当するノードもネットワーク上で近くにあることが望ましい．

以上の考察から，本論文では，特徴量空間の分割構造がネットワークノードの接続構造と同相に近くなるような手法を検討する．多くのセンサネットワークが 2 次元上に配置されるとみなせることを考慮して，SIFT 特徴量空間のある 2 次元の部分空間を考え，その部分空間内で分割を行う．

### 3.2 分割方法

分割領域をまたいだ探索の手法を提案し，評価を行う．

特徴量空間の分割領域は各ノードはあらかじめ知らないものとする．互いのノードの物理的位置は例えば GPS など各ノードは取得しているものとする．各ノードに，特徴量空間の 2 次元部分空間内の 1 点の座標を割り当てる．このとき，ネットワーク上で近くにあ



Fig. 3 特徴量空間がボロノイ型に分割された結果(ノード64)

るノードには、互いに近い座標が割り当てられるようにする。本論文では、特徴量は $600 \times 600$ の2次元部分空間の座標に射影されるので、各ノードの物理座標を相似変換させて、2次元部分空間内の1点の座標を割り当てる。

各ノードは画像から特徴点を得て、その特徴量を2次元部分空間に射影したものを計算し、宛先座標とする。その特徴量データを、宛先座標に最も近い座標を持つノードが目標ノードとなるように送信する。具体的な経路制御方法としては geographic routing<sup>10)</sup>などを用いることができる。本論文では簡単に送信ノードがリンクしているノードの中で目標ノードの座標にもっとも近い座標のノードに特徴量を送信する手法をとる。

仮にすべての特徴量データが、宛先座標に最も近い座標を持つノードまで到達したとすると、以上の方法は、各ノードに割り当てられた座標を母点として、特徴量部分空間をボロノイ分割することに相当する。実際に64分割された結果を図3に示す。点の色の違いは送信されたノードの違いである。結果的にボロノイ図状に自動的に分割される。

### 3.3 分割領域をまたいだ探索

次に、分割領域をまたいだ探索の手法について説明する。先ほどの手法で、もっとも近いノードに送信した特徴量の集合をAとする。集合Aのある1つの特徴量に注目すると、その特徴量が射影された座標から、ある一定の距離内にあるノードまでその特徴量をフラッディングする。このように送信することで、各ノードは分割数を増やしても、隣接する分割領域の特徴量を得ることができる。このフラッディングによって得られた特徴量を集合Bとする。各ノードはこれらの方法で得られる特徴量の集合A, Bを区別する。対応点を探索する場合は、Aに対応付けられる特徴量をA $\cup$ Bの中から探索する。これによって推定精度の向上が期待できるが、データ通信量も増加する。実験ではその評価を行う。

### 3.4 事前計算

本論文ではSIFT特徴量空間のある2次元部分空間を考え、その部分空間の分割を行う。この2次元部分空間に特徴量が射影される際、データ通信の負荷を分散させるために特徴量は部分空間内でより分散されることが望ましい。そこで、本論文では、主成分分析を行い、特徴量が分散される、部分空間をあらかじめ求めておく。実際に学習画像の主成分分析より得られた第1主成分と第2主成分に屋内、屋外と森のシーンの特徴量を射影した図4, 5を示す。学習画像には屋内と屋外のシーン32枚から得られた特徴点30000点を使用した。結果より、屋内と屋外のシーンにおいては、第1主成分と第2主成分の方向に十分に分散していることがわかる。しかし森のシーンにおいては、上手く分散されているとは言えない。森などのシーンにおいては別途手法を考える必要がある。

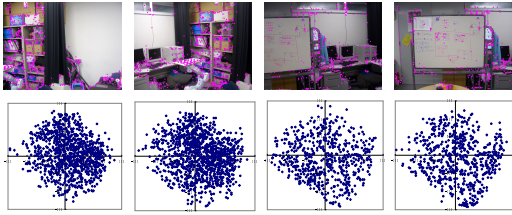


Fig. 4 サンプル画像で部分空間の射影を実行

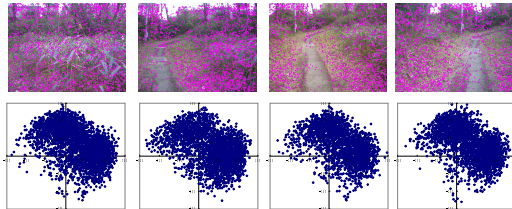


Fig. 5 サンプル画像で部分空間の射影を実行

### 3.5 視野接続推定処理

本論文においては、各ノードは先ほどの学習画像から得た主成分を取得しているものとする。この仮定下で視野接続推定を以下の手順で行う。

- 1) 各ノードに特徴量空間の2次元部分空間の1点の座標に対応づける。
- 2) 各送信ノードは特徴量データを、その特徴量との部分空間内での距離が最も近いノードに送信する(特徴量集合A)。
- 3) 各受信ノードは送信された特徴量の座標から一定距離内にあるノードへフラグgingを行う(特徴量集合B)。
- 4) 特徴量集合Aの各要素に対応づけられる特徴量をA Bの中から探索し、対応点が見つかったら、送信元のノードへレポートを返す。
- 5) 対応点の数に基づいて視野接続推定する。

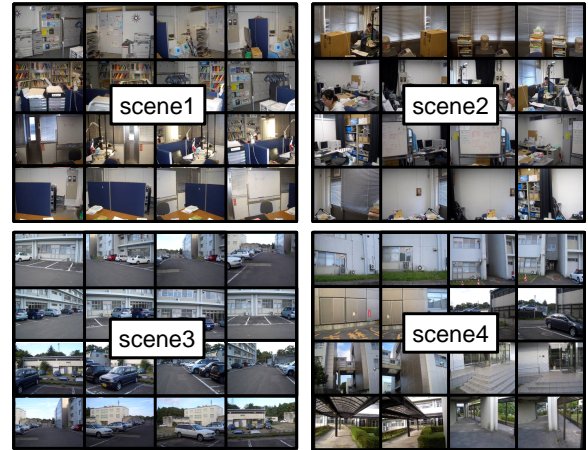


Fig. 6 使用した画像

## 4. 実験

ネットワークとして以下を仮定して実験を行う。物理的なネットワークについては、 $600 \times 600$ のフィールドにランダムにノードを配置し、ノード数ごとに定めた一定距離内のノードとネットワークを結ぶ。ノード数は16, 64, 256と変化させて評価を行う。

カメラは  $640 \times 480$  画素で、実際にはカメラ1台で図6の4シーンを16枚ずつ撮影した。視野が接続していると判断する基準は、対応点が20個以上得られた場合と経験的に定めた。2次元部分空間を定めるための学習画像としては予備評価同様のシーン1とシーン3を用いた。

推定精度の評価は、適合率 (precision) , 再現率 (recall) , F-measureを用いて行う。

式(1)は適合率を表しており、 $R$ は視野が接続と正しく推定した数、 $N$ は視野が接続と推定した数である。

$$precision = \frac{R}{N} \quad (1)$$

式(2)は再現率を表しており、 $R$ は視野が接続と正しく推定した数、 $C$ は視野が接続している正しい数である。

$$recall = \frac{R}{C} \quad (2)$$

式(3)は適合率と再現率の調和平均でありF値と呼ばれる。

$$F - measure = \frac{R}{\frac{1}{2}(N + C)} \quad (3)$$

ground truth は手動で求めた。画像中に1/8以上視野を共有していた場合に視野が接続されていると判断した。ノード数を16よりも増加させた場合の評価も行うが、ground truthの設定が煩雑になるのを避けるため、画像数はあくまで16のままとし、近隣のノードが全く同じ画像を取得するとした。

分割領域をまたいだ探索で一定の距離内のノードにフラッディングすると述べたが、この距離を変化させた時の推定精度とデータ通信量の関係を調べる。そして分割領域をまたいだ探索を行った場合、行わない場合、そしてフラッディングの場合を比較する。フラッディングは予備評価とは異なり、データが送信されていないすべてのノードに送信するという方法で行った。またクラスタリングに関しては、クラスタヘッドの周りのリンクに負荷に集中してしまうことが予備評価によりわかった点や、提案手法が分割数とノード数を等しくしている点から、クラスタリングのとの比較は行わない。

#### 4.1 領域間探索のパラメータの評価

本論文では2次元に射影された特徴量の座標は600×600のフィールドに射影している。そのフィールドのスケールで、送信する距離の半径を0～120に変化させた時のデータ通信量と推定精度の評価を行った。4つのシーンすべてにおいて同様の傾向が見られたため、シーン2のデータのみを図7, 8に示す。分割数は64とした。データ通信量は距離を増加させるほど大きくなっている。また視野接続推定精

度に関しては、F値が徐々に高くなっている。データ通信量の増加を考慮にいとると距離は60が妥当であると考え、以降の実験では距離を60にして評価を行う。

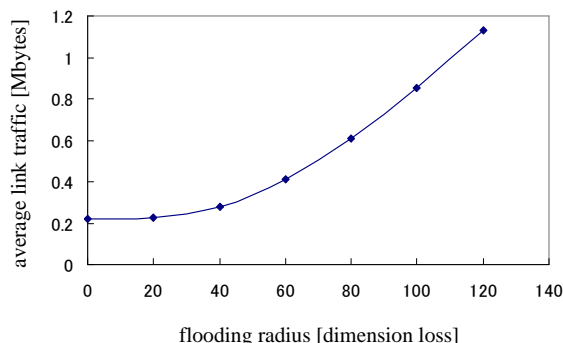


Fig. 7 フラッディングを行う距離の変化に対するデータ通信量 (シーン2, 分割数64)

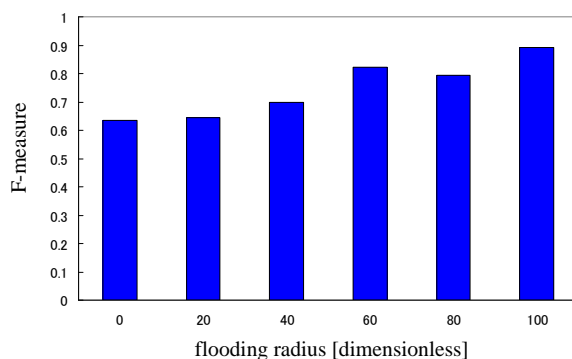


Fig. 8 フラッディングを行う距離の変化に対するの視野接続推定精度 (シーン2, 分割数64)

#### 4.2 視野接続精度の評価

視野接続推定精度の適合率と再現率の結果を図に示す。適合率における評価では、ブロードキャストと比べて提案手法は精度が落ちていくことがわかる。その原因は、提案手法において、誤った視野接続推定を多く行っているためである。また、領域をまたいだ探索を行った提案手法と行わなかった手法の結果を比較すると、領域探索を行った手法の方が適合率は高いことがわかる。また、再現率にお

ける評価では、ほぼどの手法においてよい精度を保っている。適合率が低い、つまり視野が接続と誤って推定している場合、その接続と推定したカメラ間で情報交換を行えば、実際に視野が接続しているかどうかの確認をとれる。一方で、再現率が低い、つまり視野が接続していないと誤って推定してしまった場合、実際に接続しているものを非接続と推定したままになってしまう。このことから、本論文では、適合率より、再現率の方が重要であると考えられる。ブロードキャストと提案手法を比較すると提案手法の方が適合率は低いが、再現率は同水準の精度をあげている。

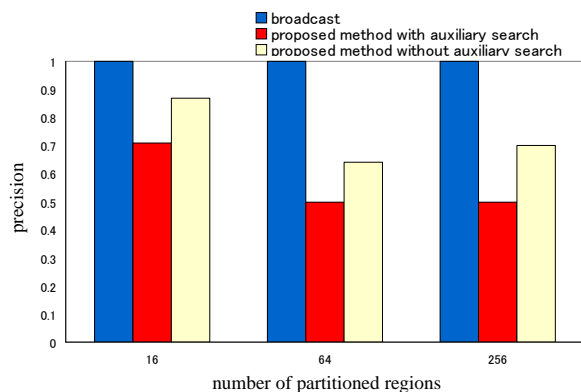


Fig. 9 適合率

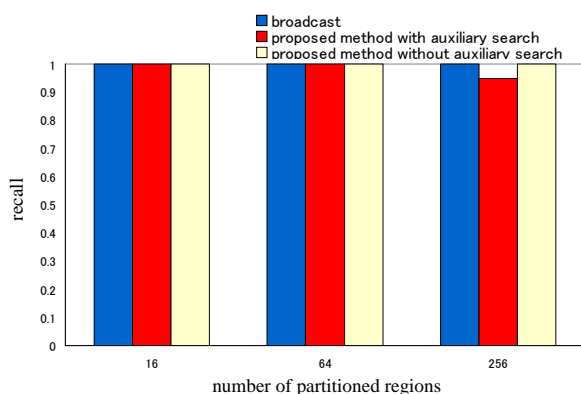


Fig. 10 再現率

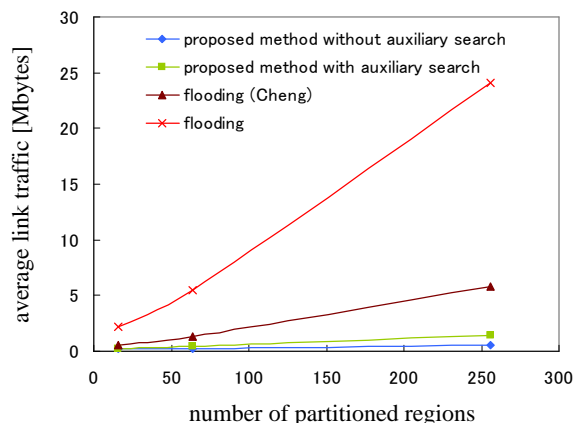


Fig. 11 データ通信量

### 4.3 データ通信量の評価

分割数の変化に対するデータ通信量の結果を、ブロードキャスト、Chengらでのブロードキャスト、領域をまたいだ探索を行った提案手法、行わなかった手法の結果を図11に示す。従来手法であるChengらの方法より大幅に低減された。

## 5. 結言

本論文では視覚センサネットワークにおいて、効率よく視野接続を推定するための手法を提案した。視野接続は画像特徴量の対応づけにより行うとし、その特徴量空間を分割することで分散化を行った。また領域をまたいだ対応点の探索も実現した。実験より、領域をまたいだ対応点の探索を行うことによって推定精度の劣化を抑えつつ、データ通信量も従来手法と比較すると大幅に低減できることがわかった。実験において、ノードの物理配置の特徴量空間へのマッピングは、各ノードは自身の物理的な位置を知っていると仮定していたため、今後は、一般のセンサネットワークで用いられる物理位置情報の取得方法<sup>11)</sup>を用いて、各ノードの位置情報を取得し、この

位置情報の正確さが結果に及ぼす影響を調べる必要がる。さらに、精度よい評価を行うために多くの画像で実験を行う必要がある。

## 参考文献

- 1) R. T. Collins, A. J. Lipton, T. Kanade, H. Fujiyoshi, D. Duggins, Y. Tsin, D. Tolliver, N. Enomoto, O. Hasegawa, P. Burt and L. Wixson, A System for Video Surveillance and Monitoring, Technical Report CMU-RI-TR-00-12, Carnegie Mellon University, 2000.
- 2) Takashi Matsuyama and Norimichi Ukita, Real-Time Multitarget Tracking by a Cooperative Distributed Vision System, *Proceedings of the IEEE*, vol. 90, no. 7, pp. 1136–1150, 2002.
- 3) Matthew Antone and Seth Teller, Scalable Extrinsic Calibration of Omni-Directional Image Networks, *Intl. J. Computer Vision*, vol. 49, no. 2/3, pp. 143–174, 2002.
- 4) Gregory C. Sharp, Sang W. Lee and David K. Wehe, Multiview Registration of 3D Scenes by Minimizing Error between Coordinate Frames, *IEEE Trans. Pattern Recognition and Machine Intelligence*, vol. 26, no. 8, pp. 1037–1050, 2004.
- 5) Henry Detmold, Anton van den Hengel, Anthony Dick, Alexc Cichowski, Rhys Hill, Ekim Kocadag, Katrina Falkner and David S. Munro: Topology Estimation for Thousand-Camera Surveillance Networks, in *First ACM/IEEE Intl. Conf. on Distributed Smart Cameras*, pp. 195–202, 2007.
- 6) Zhaolin Cheng, Dhanya Devarajan and Richard J. Radke, Determining Vision Graphs for Distributed Camera Networks Using Feature Digests, *EURASIP J. Advances in Signal Processing*, vol. 2007, Article ID 57034, 2007.
- 7) David G. Lowe, Distinctive Image Features from Scale-Invariant Keypoints, *Intl. J. Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- 8) Hari Balakrishnan, M. Frans Kaashoek, David Karger, Robert Morris and Ion Stoica, Looking Up Data in P2P Systems, *Communications of the ACM*, vol. 46, no. 2, pp.43–48, 2003.
- 9) Scott Shenker, Sylvia Ratnasamy, Brad Karp, Ramesh Govindan and Deborah Estrin, Data-Centric Storage in Sensor-nets, *ACM SIGCOMM Computer Communications Review*, vol. 33, no. 1, pp. 137–142, 2003.
- 10) Brad Karp and H. T. Kung, GPSR: Greedy Perimeter Stateless Routing for Wireless Networks, *Sixth Annual ACM/IEEE Intl. Conf. on Mobile Computing and Networking*, pp. 243–254, 2000.
- 11) David Kempe, Alin Dobra and Johannes Gehrke, Gossip-Based Computation of Aggregate Information, *44th Annual IEEE Symp. on Foundations of Computer Science*, 2003.