

森田 真司
Shinji Morita

山澤 一誠
Kazumasa Yamazawa

横矢 直和
Naokazu Yokoya

1. はじめに

近年、銀行における防犯や交差点での交通量把握、ビルのように複数の部屋や複数のビルを一括監視する場合などを目的としたカメラによる遠隔地の監視に対する要求が高まっている。監視システムの重要な機能として、環境を常に広範囲に撮影し、かつ注目すべき対象を実時間で検出・追跡することが求められる。

監視システムにおいて、従来使用されている標準レンズを備えた CCD カメラは、画角が狭く、環境全体を撮影できず、環境を広範囲に動く複数の物体の位置すべてを知ることはできない。この問題の解決策として、カメラを回転させる方法がある。しかし手動で制御する場合は、人が常に画像を注視する必要があり監視者の負担になる。画像処理により自動的に回転させる手法 [1] の場合、監視者の負担は減るが、環境全体の画像を取得できないために複数の物体の位置を同時に知ることはできない。またカメラを複数使う場合、複数の物体の位置を同時に計測できるが、環境全体を監視するには多くのカメラが必要となり、カメラ間の制御が困難である [2]。もう一つの解決策としてカメラの周囲 360° の画像が一度に取得できる全方位画像センサを用いる方法が考えられる [3, 4]。全方位画像センサを用いることにより、センサ 1 台または少数で環境全体を監視することができる。しかし従来の全方位画像センサを用いた手法は 1 台の計算機ですべてのセンサを管理しているものがほとんどで、ビル全体など多くのセンサを必要とする環境には適用できなかった。

以上の問題に対して本研究では、全方位画像センサとネットワークを利用した分散処理による実時間監視システムを開発した。本システムはサーバ/クライアントモデルであり、まずサーバ側において監視したい環境を全方位画像センサ HyperOmni Vision [5] によって撮影し、その画像より移動物体を検出し、センサからの移動物体の方位を推定する。次にその全方位画像と方位情報をネットワークを利用してクライアント側に転送する。クライアント側では時系列の方位情報を用いて移動物体の同定を行い、物体方向の画像提示による追跡を行う。これにより本システムではネットワークを介して実時間監視を行うことを可能とする。

2. 全方位画像センサ HyperOmni Vision

本研究で利用する全方位画像センサ HyperOmni Vision は、図 1 に示すように鉛直下向きに設置した双曲面ミラーとその下に鉛直上向きに設置したカメラから構成される。これにより図 2 のようにセンサの周囲 360° を一度にカメラで撮影することができる。また、入力画像は双曲面ミラーの内側焦点を投影中心とした双曲面への中心投影画像であり、入力画像を内側焦点からみた一般の平面透視投影画像図 3 やパノラマ画像に変換できる。本研究ではサーバ側で撮影した全方位動画画像をネットワークを通じてクライアント側に送り、実時間で物体方向の平

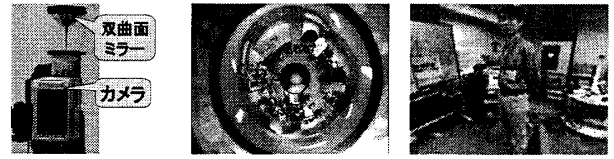


図 1: HyperOmni 図 2: 全方位画像 図 3: 平面透視投影
Vision の外観 画像

面透視投影画像に変換し監視者に提示する。

3. 移動物体の検出

本節では注目する移動物体をロバストに検出する手法について述べる。ここで、注目物体を、監視環境中に新たに現れた物体、もしくは監視環境の中で移動する物体とする。本研究では背景差分に基づいた注目物体の検出手法を採用した。まず入力画像中の背景画素の輝度 I を以下の式のように仮定する。

$$I = \bar{I} + \sigma \sin(2\pi\omega t) + k\zeta \quad (1)$$

上式は入力画像中の各背景画素について成立し、 \bar{I} は輝度値の時間平均、 σ は輝度の振幅、 ω は輝度の周波数、 t は時間、 k は $-1, 0, +1$ のいずれかの値、 ζ はカメラのみに依存した雑音を表す。また、 $\sigma \sin(2\pi\omega t)$ の項は蛍光灯や CRT ディスプレイなどのフリッカー、窓の外の樹木の揺らぎなどを表し、 $k\zeta$ の項はゲインアップによるごま塩ノイズなどカメラにのみ依存する雑音を表す。このとき背景画素の輝度値 I は以下の範囲をとりうる。

$$\bar{I} - \sigma - \zeta \leq I \leq \bar{I} + \sigma + \zeta \quad (2)$$

本手法では、入力画素の輝度 I が式 (2) の範囲におさまる場合はその画素は背景の画素、そうでない場合は物体の画素とする。

あらかじめ移動物体が存在しない状況で ζ を測定しておく。この ζ はカメラのみに依存する値であるので、環境によって変更する必要はない。背景のゆっくりとした変化を考慮し、各画素の輝度 I が式 (2) の範囲におさまる背景であると判断した場合、式 (3) を用いて毎フレームごとに \bar{I} と σ を更新する。ただし、 \bar{I}' と σ' は次フレームの \bar{I} と σ とする。

$$\begin{aligned} \bar{I}' &= (n-1)/n \times \bar{I} + 1/n \times I \\ \sigma' &= (n-1)/n \times \sigma + 1/n \times 2 \times (I - \bar{I})^2 \end{aligned} \quad (3)$$

n: 更新速度パラメータ

また、各画素が物体の部分であると判断された場合でも、急激な背景の変化や新たに監視環境中に置かれた静止物体に対応するため、式 (4) を用いて σ を更新する。

$$\begin{aligned} \bar{I}' &= \bar{I} \\ \sigma' &= (m-1)/m \times \sigma + 1/m \times 2 \times (I - \bar{I})^2 \end{aligned} \quad (4)$$

m: 更新速度パラメータ ($m \geq n$)

これによって、新たな物体などがその場に存在し続けた場合には σ が増加し、しばらくするとその物体が背景と判断されるようになる。

次に注目物体の方位情報を求める。具体的には毎フレームごとに物体とみなされた画素から図 4 に示すよう

な θ 方向の画素数のヒストグラムを作成する。このヒストグラムから連続した 0 より大きい値を持つ部分を注目物体の方位角範囲とする。さらに、各方位角範囲で検出された物体を囲むように仰角の範囲も求め、注目物体の方位情報とする。

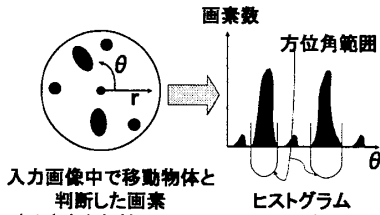


図 4: 極座標変換によるヒストグラムの作成

4. サーバとクライアント間の通信

本システムは、監視者側に設置した PC をクライアント、監視環境に設置した PC をサーバとしたサーバ/クライアントモデルであり、ネットワークを用いて二者間のデータ通信を行うことで分散処理を行い、ネットワークを介した遠隔監視を実現する。また、本システムでは Digital Video Transport System (DVTS) [6] を用いてネットワークで全方位動画を伝送する。DVTS はネットワークを利用してデジタルビデオ画像 (DV データ) を伝送できるシステムであり、33Mbps の帯域幅を使用して高品質動画 (640 × 480, 30fps) を伝送することが可能である。

4.1 サーバ側の処理

サーバ側では全方位画像センサにより撮像された全方位画像から 3 節の手法により複数の注目物体の方位情報を計算する。また、デーモンプロセスにより、クライアントの要求に基づいてコネクションを確立し、全方位動画と複数の注目物体の方位情報を送信する。

4.2 クライアント側の処理

クライアント側ではサーバとのコネクション確立後、全方位動画と複数の注目物体の方位情報を受信し、時系列での複数の物体の同定を行い、監視者に物体方向の画像を提示する。

5. 注目物体の追跡

本節では、クライアントにおいて時系列での注目物体の同定を行い、監視者に物体方向の画像を提示する手法について述べる。サーバから得られた複数の注目物体の方位情報から時系列で方位の変化が最も小さくなるように複数の注目物体の方位の対応付けを行い、注目物体の同定を行う。最後に受信した全方位動画から同定された複数の物体方向の平面透視投影画像を生成して監視者に提示する。全方位動画から平面透視投影画像へは参考文献 [7] の手法によりハードウェアのテクスチャマッピング機能を用いて実時間で変換できる。

6. システム構成と実験

図 5 のようなシステム構成で、本学情報科学研究科棟のロビー (1 階) を監視するサーバシステムを置き、研究室 (3 階) を監視者側としてクライアントシステムを

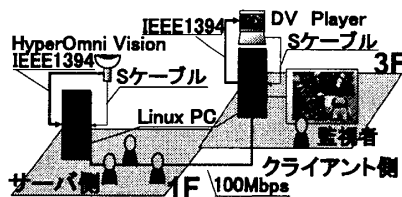


図 5: システム構成

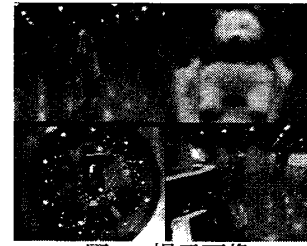


図 6: 提示画像

置いて、学内 LAN を用いて動作実験を行った。監視環境中に複数の人物が歩行し、さらに 1 つの物体を環境中に新たに置いた状態での監視を試みた。なお、サーバ、クライアント側とも、計算機は PC (Pentium IV 2.0GHz, メモリ 512MB, Linux) 各 1 台であり、ネットワークは有線 100Mbps である。

クライアント側で監視者に提示された画面の例を図 6 に示す。本実験では画面を 4 分割し、左下に全方位画像を、残り 3 つに検出された物体を入力画像中で面積の大きい方から提示した。図 6 において 2 人の人物と新たに置かれた物体が検出できていることがわかる。サーバ側における物体検出により方位情報は約 0.2 秒間隔で更新できた。またクライアント側においては平面透視投影画像の方位は約 0.2 秒ごとに更新されるものの、監視者に提示される画像は 22fps で更新できた。また蛍光灯などのフリッカーなどは無視し、ロバストに移動物体のみを検出できることを確認した。

7. まとめと今後の課題

本報告では、全方位画像センサによるネットワークを介した遠隔監視システムについて述べた。本システムは 1 対 1 のサーバ/クライアントシステムであり、ネットワークを利用して分散処理を行った。背景差分を利用し、かつ背景の更新を行うことで、ロバストに複数の移動物体を検出し、クライアント側において平面透視投影画像を提示することにより移動物体の確認が可能であった。今後の課題としては、複数のサーバとクライアントを配置した多対多のシステムへの拡張が挙げられる。また、監視者の使用方針に基づいた注目すべき提示画像の選択や全方位画像センサの選択方法の検討が挙げられる。

参考文献

- [1] S. Rougeaux and Y. Kuniyoshi: "Velocity and disparity cue for robust real-time binocular tracking," Proc. of CVPR'97, pp.1-6, 1997.
- [2] T. Kanade, T. Collins, A.J. Lipton, P. Anandan, P. Burt and L. Wixson: "Cooperative multi-sensor video surveillance," DARPA Image Understanding Workshop, vol.1, pp.115-122, 1998.
- [3] 十河, 石黒, Trivedi: "複数の全方位画像センサによる実時間人間追跡システム," 信学論, Vol.J83-D-II, No.12, pp.2567-2577, 2000.
- [4] 寺沢, 山澤, 竹村, 横矢: "複数の全方位画像センサを用いた遠隔監視システムにおける複数移動物体の存在領域推定," 信学技報, PRMU2000-195, 2001.
- [5] 山澤, 八木, 谷内田: "移動ロボットのための全方位視覚センサ HyperOmni Vision の提案," 信学論, Vol.J79-D-II, No.5, pp.698-707, 1996.
- [6] A. Ogawa, K. Kobayashi, O. Nakamura, and J. Murai: "Design and Implementation of DV Stream Over Internet," Proc. IWS Internet Workshop, No.99EX385, 1999.
- [7] Y. Onoe, K. Yamazawa, H. Takemura, and N. Yokoya: "Telepresence by real-time view-dependent image generation from omnidirectional video streams," Computer Vision and Image Understanding, Vol.71, No.2, pp.399-406, 1992.