

画像理解システム評価のための多視点動画画像データベースの開発について

浮田, 宗伯
京都大学大学院情報学研究科

東海, 彰吾
京都大学大学院情報学研究科

松山, 隆司
京都大学大学院情報学研究科

谷口, 倫一郎
九州大学大学院システム情報科学研究科

<https://hdl.handle.net/2324/5790>

出版情報：電子情報通信学会技術研究報告. PRMU, パターン認識・メディア理解. 99 (182), pp.65-72, 1999-07-16. 電子情報通信学会

バージョン：

権利関係：

画像理解システム評価のための 多視点動画像データベースの開発について

浮田宗伯† 東海彰吾† 松山隆司† 谷口倫一郎‡

† 京都大学大学院情報学研究科, 京都市左京区吉田本町

‡ 九州大学大学院システム情報科学研究科, 福岡市東区箱崎 6-10-1

著者らは、画像理解システムの評価環境を整備することを目的として、標準的な画像データセットの作成を進めている。このような画像データベースは、客観的、科学的立場から各種画像理解アルゴリズムの性能評価を行うための研究基盤として有効であると考えられる。今回は、最近注目されている、多視点動画像を用いた動的シーン解析のための画像データセットの作成を行った。ここでは、撮影に利用したシステムおよび、撮影した画像セットの概要について報告する。

キーワード: 画像理解システム評価, 多視点動画像, 動画像データベース

Database Development of Multi-Viewpoint Image Sequences for Evaluation of Image Understanding Systems

Norimichi UKITA† Shogo TOKAI† Takashi MATSUYAMA†
Rin-ichiro TANIGUCHI‡

† Graduate School of Informatics, KYOTO University

Yoshidahonmachi, Sakyo, Kyoto, Japan

‡ Graduate School of Information Science and Electrical Engineering, KYUSHU University

6-10-1, Hakozaki, Higashi, Hukuoka, Japan

We are developing a set of standard images for comparative evaluation of image understanding algorithms and systems. This paper describes multi-viewpoint image sequence database which we are developing. Following the specification of the newly developed image capturing system, we show several illustrative examples of captured image sequences.

Keywords: Image Understanding System Evaluation, Multi-Viewpoint Image Sequence, Dynamic Image Database

1 はじめに

これまでの画像理解研究によって、数多くの知見やアルゴリズムが得られている。しかし、これらの中で、どのような問題が、どのような環境設定の下で、どこまで解決され、何が未解決であるか、また、他の研究と比較した際の優劣を明確に述べることができるものは少ない。この問題を残す限り、過去の研究成果の上に新たな成果を積み上げていくことは難しい。その原因として、以下の二つのものが考えられる。

問題 1: アルゴリズムの有効性の検証法の欠如

問題 2: 対象世界(問題)の複雑さの規定基準の欠如

我々は、共通のデータを用いた比較実験を行うことにより、アルゴリズムの有効性を評価することを目的として画像データベースの開発を行っている [1][2]。これには、以下のような意義、目的がある。

1. センサ、光源、撮影対象物体などの情報を求めておく。これにより、画像解析結果を客観的、定量的に評価可能となり、問題 1 の解決に役立つ。
2. 対象世界や認識対象の持つ複雑さを表す種々の尺度を選定し、各尺度に関して数種類の具体的な画像を撮影、データベース化する。これにより、上述の問題 2 の解決に向けての基盤が整備される。

これまで、平成 6~8 年度にかけて、各種画像理解アルゴリズムの性能評価を客観的、科学的立場から行うための研究基盤整備を目的として、キャリブレーションデータ付き画像の収集・作成を行った [1][2]。現在は、動的シーンを対象とした多視点動画データ作成、収集を目指した実時間動画入力法の検討と実験的な撮影を行っている。多視点動画を撮影する際の問題点には、以下のものが挙げられる。

問題 A: 各種画像理解アルゴリズムの性能評価を行うのに十分な精度の実時間動画撮影法

問題 B: 複数カメラの同期撮影法

問題 C: 各カメラのキャリブレーション

現在既に、文献 [3] のような多視点動画データベースが存在するが、このデータベースの仕様は、

a: VTR を介して A/D 変換された数秒間の動画像 (lossless 圧縮画像, MPEG2, MPEG1 の 3 種類)

b: 1 台のカメラで発生したタイムコードを他のカメラに入力しておき、A/D 変換時に全ビデオのタイムコードを合せることによりカメラ間同期を実現

c: カメラキャリブレーション用画像が付属しているであり、中間記録媒体の使用、厳密でないカメラ間同期など、上記問題 A,B,C を同時に解決するのに十分なデータではない。そこで我々は、この問題解決のための撮影方法・システムを考案・開発し、動画データの作成を行った。また、各データには各種キャリブレーションデータを添付している。主な成果物を以下に示す。

局所空間における複数運動対象の多視点実時間動画

- 両手を使った作業の 6 視点実時間動画像
- 手話の 6 視点実時間動画像

室内空間における複数運動対象の多視点実時間動画

室内における人体の全体的動きの 6 視点動画像

- 1 人による体操シーン
- 2 人による格闘シーン

人物の全方位ステレオ画像

人物大の静止対象を、キャリブレーションされた 9 台のカメラによって、全方位視点撮影画像

屋外空間における複数運動対象の多視点実時間動画

屋外における車両の広域移動の 2 視点動画像

2 画像撮影システム

本章では、今回開発した 3 種類の画像撮影システムの詳細について述べる。

2.1 同期型多視点動画撮影システム

本システムは、以下の目的を実現するための多視点動画撮影システムである。

1. 全カメラの同期駆動
2. カメラの最大撮像能力で得られる画像を非圧縮記録
3. VTR などの中間媒体を介さないデジタル化
4. カメラの台数を任意に調節可能

1. の実現のために、全カメラは同一の外部同期信号によって駆動している。また、2,3,4. の実現のために、カメラと同じ台数のネットワーク接続された PC にキャプチャボードを挿入し、並列に画像を取り込んでいる。

この方式を採用した際の問題は、各 PC での取り込みタイミングを合わせる必要があるという点である。今回の撮影では、各 PC の内部時計を NTP によって同期させ、画像を取り込んだ時刻を調べることにより、各 PC での取り込みタイミングの同定を行った。NTP を用いれば、LAN 上の PC については、数ミリ秒の精度で時計が合わせられることが知られており、画像取り込み間隔である 30msec に比べて十分小さいため、画像取り込みタイミングの同定には十分である。

今回実装したシステムでは、高速なディスク装置がまだ高価であるため、取り込んだ動画を主記憶上のバッファに一度書き込んだ後、オフラインでディスクに書き出している。そのため、一回に撮影できるシーンは 640×480 のカラー画像で、数十秒程度までである。以下に、使用した機器等を示す。

センサ:

センサ本体: Victor KY-F57

レンズ:

SONY VCL-16WM 16mm(データ bomb, sign)

FUJINON TF-4DA-8 4mm(データ ciko, figh)

センサパラメータ:

絞り: マニュアル

フォーカス値: マニュアル

シャッター速度: 1/60 sec (NORMAL)

補正:

γ 補正: 無
Green CCD 位置ずれ: 無
エッジ強調: 無

中間記録: 無

インターレース: 有 (偶数フィールド)

外部同期発生装置: Tektronix TSG-200

キャプチャボード: Imaging Technology ICPCI-RGB

入力信号: アナログ RGB

計算機: PCI コントローラのチップセットが Intel
440LX 以降のマザーボードを搭載した PC

OS: Linux kernel 2.0.36

デバイスドライバ: ICPCI Driver for Linux から開発¹

NTP: NTP Ver.3-5.93

2.2 長基線長多視点静止画像撮影システム

本システムは、静的シーンを対象として、物体をあらゆる方向から撮影することを目的とした静止画像撮影システムである。各カメラの映像信号は、一対一接続された PC のキャプチャボードにそれぞれ入力される。

ここで使用したカメラは、全てほぼ視点固定型と仮定できる首振りカメラであり、PC から、パン角 (-30° ~ 30°), チルト角 (-15° ~ 15°) などのパラメータを制御可能である。以下に、使用した機器等を示す。

センサ:

センサ本体: SONY EVI-G20

レンズ: 不明 (本体に装着のもの)

センサパラメータ:

ズーム値 (焦点距離): 最広角側に設定

絞り: マニュアル

フォーカス値: マニュアル

シャッター速度: 1/60 sec

補正:

 γ 補正: 有 (γ 値不明)

Green CCD 位置ずれ: 一板式

エッジ強調: 有

中間記録: 無

インターレース: 有 (奇偶両フィールド)

キャプチャボード: Matrox METEOR RGB

入力信号: コンポジット信号

計算機: Pentium 166MHz 搭載 PC

2.3 同期型 2 視点交通動画撮影システム

本システムは、以下の目的を実現するための多視点動画撮影システムである。

1. 全カメラの同期駆動
2. 屋外の対象を撮影
3. 記録時間の長い動画の撮影
4. カメラの台数を任意に調節可能

1. の実現のために、全カメラは同一の外部同期信号によって駆動している。また、2.3.4. の実現のために、各カメラと一対一に接続されたビデオデッキを用いて撮影対象シーンを一旦ビデオテープに録画した後、それを再生しながら PC へ取り込んでいる。

¹ オリジナル版共々 <http://www.gom-online.de> から入手可能

この方法を採用した場合、PC へ画像を取込む際、ビデオテープ録画時に同期のとれているフレームを確認しながら取り込みを行う必要がある。ここでは、ビデオテープ録画時に、タイムコードをデッキ 1 からデッキ 2 に分配することにより、2 台のデッキで同一のタイムコードを同時記録しておき、PC への取り込みの際にそのタイムコードを 2 つのビデオテープ間で合わせることにより、取込み画像の同期を保証している。

今回撮影したデータセットの記録時間は、1 分、または 1 分 30 秒であるが、一旦ビデオテープに記録した画像を PC へ取込む方法を採用しているため、もっと長時間にわたるシーンの変化を撮影した動画の作成も容易である。以下に、使用した機器等を示す。

センサ:

センサ本体: Victor KY-F57

レンズ: Cannon PH6X8 MACRO

センサパラメータ:

ズーム値 (焦点距離): 最広角側に設定

絞り: マニュアル

フォーカス値: マニュアル

シャッター速度: 1/60 sec

EEI が ON の場合はシャッター速度が変化。

補正:

 γ 補正: データによって有 (γ 値不明), 無

Green CCD 位置ずれ: 無

エッジ強調: データによって不明, 無

明度補正: データによって有, 無

ALC (ゲイン調整) と EEI (シャッター速度制御) の 2 種類の明度補正がある。

中間記録: 有 (S-VHS, S-ビデオ信号)

インターレース: 有 (奇偶両フィールド)

キャプチャボード: SGI O2 内蔵ビデオ I/F

入力信号: S-ビデオ信号

計算機: SGI O2

3 データベースの仕様

本章では、今回撮影した各種動画データの詳細について述べる。

3.1 同期型多視点動画

3.1.1 共通事項

ここでは、節 3.1 で紹介するデータセットに関する共通事項について述べる。

想定する問題の記述: 多視点動画を用いた動物体追跡あるいは動作の解析

画像の種別:

時間変化: 動画

センサ台数: 6 台, 備考: 同期駆動している。

スペクトル: RGB カラー

画像サイズ: 320 × 240 (RGB 各 8bit)

備考: インターレース方式の影響を除去するため、640 × 480 で獲得した画像フレームのフィールドを画像として使い、320 × 240 にリサイズしている (水平方向のみリサイズ)。

画像ファイルフォーマット: PPM

備考:画像間の同期保証のため,各画像のコメント部分にNTPのタイムスタンプを付加している.

エネルギー源:

エネルギー源のタイプ:

種別: 室内照明
 照明の種類: 高周波蛍光灯
 ワット数: 40W
 個数: 12本
 照明領域の面積(床面積): 5.2m × 6m
 シーンでの距離: 約2m

撮影対象物体(人間):

光学的特性:

テクスチャの有無: 有
 表面反射の強弱: 弱
 シェーディングの有無: 無
 影の投影の有無: 無(データbomb,sign),小(データciko,figh)
 主な領域の色: 肌色,黒,白,赤
 主な領域の明るさ(暗1~5明): 3(データbomb,ciko,figh),2(データsign)

動的特性: 多関節物体(時間的変化有)

3.1.2 局所空間における複数運動対象の多視点実時間動画: bomb

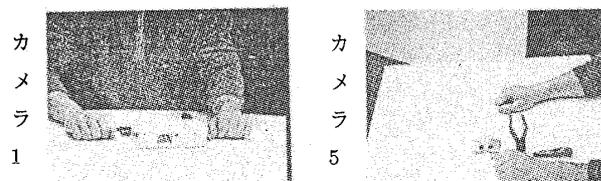


図1: bomb: サンプル画像(初期フレーム)

撮影対象となったシーンの状況は,図1のように,ドライバとベンチを用いて配線を切るシーンである.このシーンを,500フレーム×6視点計3,000枚だけ撮影した.撮影時の,物体,センサの配置を図2に示す.

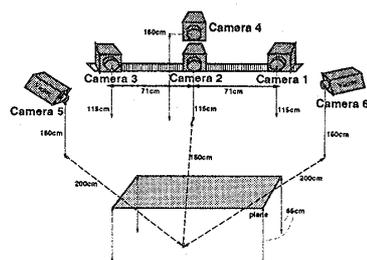


図2: 配置図(センサ間の距離はおよその値)

撮影対象となった物体に関するデータを以下に示す.

対象物体(箱):

光学的特性:

テクスチャの有無: 有
 表面反射の強弱: 弱
 シェーディングの有無: 無
 影の投影の有無: 無
 主な領域の色: 白
 主な領域の明るさ(暗1~5明): 5

動的特性: 剛体(時間的変化有)

ドライバ,ベンチ:

光学的特性:

テクスチャの有無: 有
 表面反射の強弱: 強
 シェーディングの有無: 無
 影の投影の有無: 無
 主な領域の色: 青,銀
 主な領域の明るさ(暗1~5明): 2

動的特性: 剛体(時間的変化有)

パネル:

光学的特性:

テクスチャの有無: 無
 表面反射の強弱: 弱
 シェーディングの有無: 無
 影の投影の有無: 無
 主な領域の色: 白
 主な領域の明るさ(暗1~5明): 5

動的特性: 剛体(時間的変化無)

3.1.3 局所空間における複数運動対象の多視点実時間動画: sign

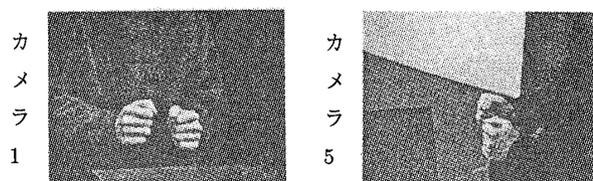


図3: sign: サンプル画像(初期フレーム)

撮影対象となったシーンの状況は,図3のように,1人で手話するシーンである²このシーンを,500フレーム×6視点計3,000枚だけ撮影した.撮影時の,物体,センサの配置を図2に示す.

撮影対象となった物体に関するデータを以下に示す.

台:

光学的特性:

テクスチャの有無: 無
 表面反射の強弱: 弱
 シェーディングの有無: 無
 影の投影の有無: 無
 主な領域の色: 黒
 主な領域の明るさ(暗1~5明): 5

動的特性: 剛体(時間的変化無)

3.1.4 室内空間における複数運動対象の多視点実時間動画: ciko

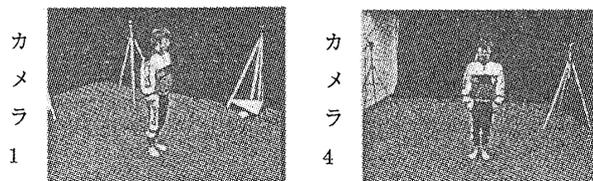


図4: ciko: サンプル画像(初期フレーム)

撮影対象となったシーンの状況は,図4のように,屈伸運動³などを連続的に行うシーンである.このシーンを,1000フレーム×6視点計6,000枚だけ撮影した.撮影時の,物体,センサの配置を図5に示す.

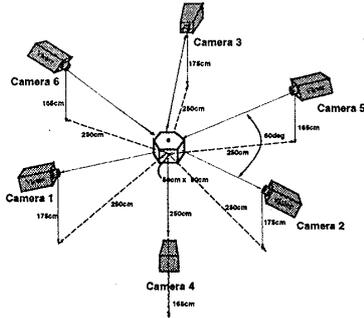


図 5: 配置図 (センサ間の距離はおよその値)

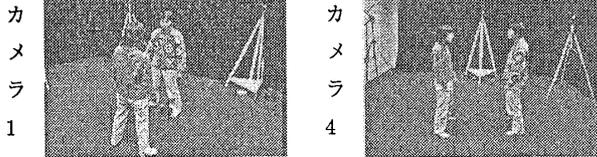


図 6: fig: サンプル画像 (初期フレーム)

3.1.5 室内空間における複数運動対象の多視点実時間動画: fig

撮影対象となったシーンの状況は、図 6 のように、二人で格闘するシーンである。このシーンを、1000 フレーム×6 視点計 6,000 枚だけ撮影した。撮影時の、物体、センサの配置を図 5 に示す。

3.2 長基線長 9 視点静止画像: 9cam

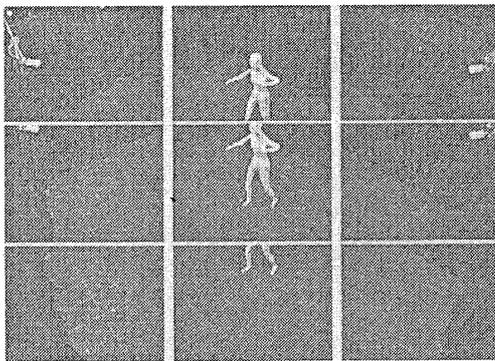


図 7: 9cam: サンプル画像

想定する問題の記述:

- 多視点画像を用いた形状復元。
- 多視点撮影システムの較正法の評価。

画像の種類:

時間変化: 静止画
 センサ台数: 複数 (9 台)
 スペクトル: RGB カラー
 画像枚数:

$$9[\text{camera}] \times 9[\text{image}] \times 8[\text{variation}] = 648[\text{image}]$$

画像サイズ: 640 × 480 (RGB 各 8bit)

画像ファイルフォーマット: PPM

物体、光源、センサの配置図: 図 8

備考: センサ間の距離はおよその値を示している。また、図中の座標系は、添付したカメラの外部パラメータとは異なる座標系であることに注意。

² 手話 (と指文字) の意味は「今日は楽しかったです。あなたの声しか届きません。今度一緒に遊びに行きましょう。お願いします。」。

³ 筑高体操と呼ばれる、福岡県立筑紫丘高等学校の体操である。

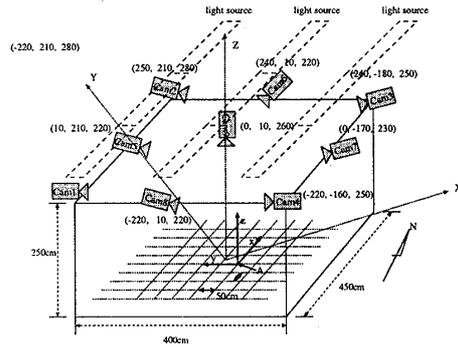


図 8: 物体、センサの配置図

エネルギー源:

エネルギー源のタイプ:

種別: 室内照明
 照明の種類: 蛍光灯 (色不明)
 ワット数: 50 ワット
 個数: 18 本

撮影対象となったシーンは、図 7 のように、床に目印として 50cm 毎に格子状に糸を貼った室内である。ここに、各種対象物体を設置して計 8 種類の画像セットを撮影した。以下に、各シーン状況の説明を示す。

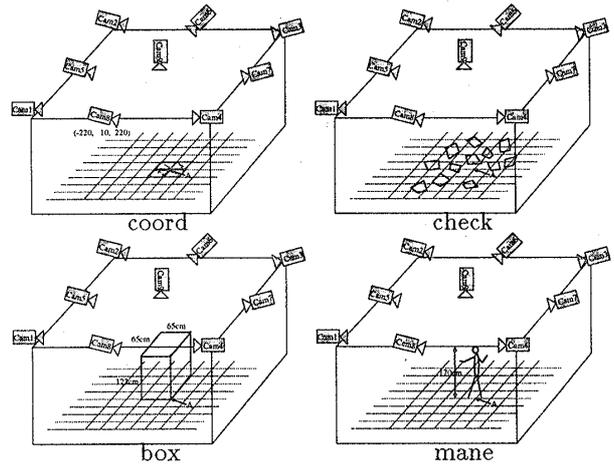


図 9: シーンの概略図 (9cam)

- 状況 1 - 物体無し (back):
対象物体: 無
- 状況 2 - キャリブレーション用 (coord):
対象物体: 数字 (1~4) を刷った紙 (4 枚) を床に配置
シーンの概略図: 図 9(coord).
- 状況 3 - キャリブレーション用 2 (check):
対象物体: 数字, 英字, 及び格子模様を印刷した紙 (多数) を床に配置
シーンの概略図: 図 9(check).
- 状況 4 - 箱 (box):
対象物体: およその大きさが既知の直方体の箱をほぼ中央に設置。
シーンの概略図: 図 9(box).
- 状況 5 - マネキン 1 (mane1):
対象物体: マネキンをほぼ中央に設置。座姿勢
シーンの概略図: 図 9(mane).

● 状況 6 - マネキン 2 (mane5):

対象物体: マネキンをほぼ中央に設置. 立姿勢
シーンの概略図: 図 9 (mane).

● 状況 7 - 着衣マネキン 1 (mane3):

対象物体: 着衣マネキンをほぼ中央に設置. 立姿勢
シーンの概略図: 図 9 (mane).

● 状況 8 - 着衣マネキン 2 (mane4):

対象物体: 着衣マネキンをほぼ中央に設置. 座姿勢
シーンの概略図: 図 9 (mane).

対象となった物体に関するデータを以下に示す.

箱:

光学的特性:

テクスチャの有無: 有
表面反射の強弱: 弱
シェーディングの有無: 有 (微少)
影の投影の有無: 無
主な領域の色: 茶
主な領域の明るさ (暗 1~5 明): 3

幾何学的形状: 直方体 (65cm × 65cm × 123cm)

動的特性: 剛体

マネキン:

光学的特性:

テクスチャの有無: 無 (服無), 有 (服有)
表面反射の強弱: 強 (服無), 弱 (服有)
シェーディングの有無: 有
影の投影の有無: 無
主な領域の色:

白 (服無), 桃色 (服), 青 (ズボン) 灰色 (帽子)

主な領域の明るさ (暗 1~5 明):

4 (服無), 2 (服有)

幾何学的形状: 人型 (身長 170cm)

動的特性: 多関節物体

3.2.1 キャリブレーション

較正用に用いることができるデータ: 較正に使用できるデータとしては, 以下の画像などが使用できる.

- 床面に張った 50cm の格子模様
- 床面に数字, 英字, 格子模様を印刷した紙を並べた画像 (check)
- サイズ既知の箱を撮影した画像 (box)
- 白いスクリーンになるべく均等に光を当て撮影した画像 (口径食補正用)

内部パラメータの較正: 視点固定型首振りカメラのための内部パラメータキャリブレーション法により, 上のデータ画像から撮影に用いたカメラの内部パラメータを求めた [4]. 決定したパラメータの例を以下に示す.

-8.760387e-08 歪み係数
332.985021 歪み中心の x 座標 [pixel]
282.044026 歪み中心の y 座標 [pixel]
0.000000 回転軸と投影中心のずれ (未使用)
0.000000 回転軸と投影中心のずれ (未使用)
-0.256260 CCD 面の光軸周りの回転 [degree]
825.022033 焦点距離
0.561196 CCD 面のパン角 [degree]
-0.115157 CCD 面のチルト角 [degree]
0.998137 アスペクト比

外部パラメータの較正: 全撮影画像中に観測されている平面領域中のテクスチャの見え方が, 各画像間どのように変化しているかを解析することにより, 各カメラの外部パラメータを弱較正可能である. データベースに添付した外部パラメータは, カメラ 9 から床面に下ろした垂線を z 軸, その足を原点とし, 床面に沿ってカメラ 3 から床面に下ろした垂線の足へ向かう方向を x 軸, x 軸 z 軸に直交するように y 軸を取っている. また, カメラ 1 の投影中心からカメラ 3 の投影中心までの距離を 500 としてスケールしている. このようにして決めた外部パラメータの例を下に示す.

-0.08582186	-0.66788143	0.73930305	-314.566684
-0.99631007	0.05821807	-0.06306266	11.6920629
-0.00092242	-0.74198723	-0.67041338	314.013675
0.00000000	0.00000000	0.00000000	1.00000000

この行列を T とすると, 空間中のある点の世界座標系での座標値 (X, Y, Z) とカメラ座標系での座標値 (x, y, z) は, 以下の関係を持つ.

$$(X \ Y \ Z \ 1)^T = T (x \ y \ z \ 1)^T \quad (1)$$

3.3 2 視点同期撮影道路状況動画画像

3.3.1 共通事項

ここでは, 節 3.3 で紹介するデータセットに関する共通事項について述べる.

想定する問題の記述:

- 屋外での多視点動画画像を用いた動物体追跡.
- 交通状況認識

画像の種別:

時間変化: 動画
センサ台数: 2 台, 備考: 同期駆動している.
スペクトル: モノクロ
画像サイズ: 720 × 486 (8bit モノクロ)
画像ファイルフォーマット: PGM

エネルギー源:

エネルギー源のタイプ:
種別: 太陽光
天候: 晴天, 曇天 (1 データ内での変動有)
撮影時間: 14 時 ~ 16 時

対象物体:

道路:

光学的特性:

テクスチャの有無: 無
表面反射の強弱: 弱
シェーディングの有無: 無
影の投影の有無: 有
主な領域の色: 灰色
主な領域の明るさ (暗 1~5 明): 2

動的特性: 剛体

車両:

光学的特性:

テクスチャの有無: 不明
表面反射の強弱: 強
シェーディングの有無: 有 (少)

影の投影の有無: 有
 主な領域の色: 不明
 主な領域の明るさ(暗1~5明): 4
 動的特性: 剛体

3.3.2 2視点同期撮影道路状況動画像
 (KyotoHotel1)

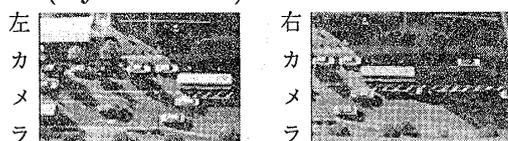


図 10: KyotoHotel1: サンプル画像

カメラは京都ホテル17階に設置し、真下の道路(河原町通り)の交通状況を撮影した(図11)。撮影は、道路が画像走査線にほぼ並行に観測されるように、且つ、互いの観測画像間で若干の重なりを持つように行い(図12)、照明変動、道路状況を変化させて、8通りのデータセットを撮影した(機器設定ALC, EELは全データセットにおいてONに固定)。各データセットは1800フレーム×2視点計3,600枚の画像からなっている。

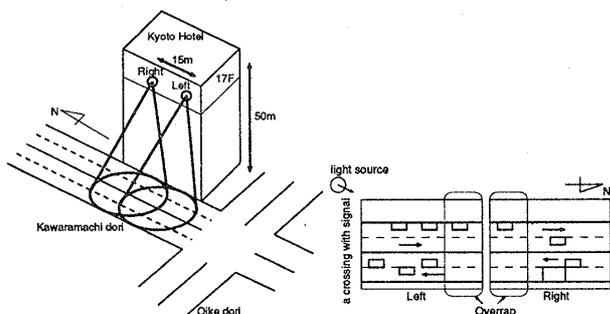


図 11: 物体、センサの配置図 図 12: シーンの概略図

- セット 1: 照明変動: 有(晴天→曇天)
道路状況: 通行量少
- セット 2: 照明変動: 無(曇天)
道路状況: 通行量少, 右折車線待車両多, 路肩から走行車線への侵入車両有
- セット 3: 照明変動: 無(曇天)
道路状況: 通行量少
- セット 4: 照明変動: 無(曇天)
道路状況: 通行量少, 大型車両多, 路肩から走行車線への侵入車両有
- セット 5: 照明変動: 有
道路状況: 通行量多, 右折車線待車両多,
- セット 6: 照明変動: 無(曇天)
道路状況: 通行量多, 右折車線待車両多, 車両の流れ遅い
- セット 7: 照明変動: 無(曇天)
道路状況: 通行量少, 右折車線待車両多, 路肩から走行車線への侵入車両有, 走行車両の路肩停止有
- セット 8: 照明変動: 有(晴天→曇天)
道路状況: 通行量少

3.3.3 2視点同期撮影道路状況動画像
 (KyotoHotel2)



図 13: KyotoHotel2: サンプル画像

カメラは京都ホテル17階に設置し、真下の道路(河原町通り)の交通状況を撮影した(図14)。撮影は、道路が画像走査線方向に対して少し傾いて観測されるように、且つ、互いの観測画像間で若干の重なりを持つように行い(図15)、照明変動、道路状況を変化させて、4通りのデータセットを撮影した(機器設定ALC, EELは全データセットにおいてそれぞれON, OFFに固定)。各データセットは1800フレーム×2視点計3,600枚の画像からなっている。

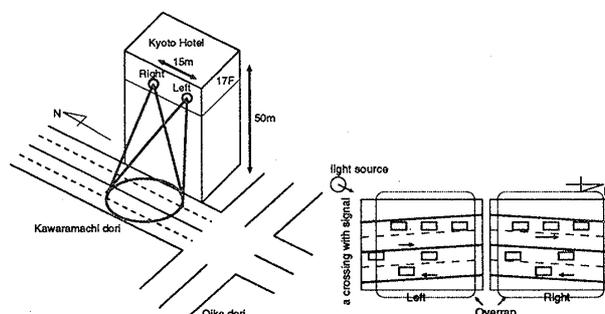


図 14: 物体、センサの配置図 図 15: シーンの概略図

- セット 1: 照明変動: 無(曇天)
道路状況: 通行量少
- セット 2: 照明変動: 無(曇天)
道路状況: 通行量多, 右折車線待車両多
- セット 3: 照明変動: 無(晴天)
道路状況: 通行量多, 右折車線待車両多
- セット 4: 照明変動: 有
道路状況: 通行量少, 右折車線待車両多

3.3.4 2視点同期撮影道路状況動画像
 (KawaramachiOike)



図 16: KawaramachiOike: サンプル画像

カメラは京都ホテル17階に設置し、道路(河原町通り)の交通状況を斜め上方から撮影した(図17)。撮影は、2台のカメラがそれぞれ、交差点とそこから続く直線道路を観測するように、且つ、互いの観測画像間で若干の重なりを持つように行い(図18)、照明変動、道路状況を変化させて、4通りのデータセットを撮影した

(機器設定 ALC, EEL は全データセットにおいてそれぞれ ON, OFF に固定). 各データセットは 1800 フレーム × 2 視点計 3,600 枚の画像からなっている.

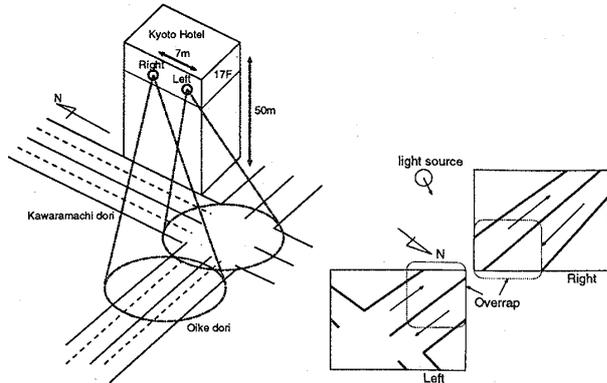


図 17: 物体, センサの配置図 図 18: シーンの概略図

- セット 1: 照明変動: 無(曇天)
道路状況: 通行量多, 遮蔽無
- セット 2: 照明変動: 無(曇天)
道路状況: 通行量少, 遮蔽無,
- セット 3: 照明変動: 無(曇天)
道路状況: 通行量多, 対向車による遮蔽有
- セット 4: 照明変動: 無(曇天)
道路状況: 通行量多, 対向車による遮蔽有

3.3.5 2 視点同期撮影道路状況動画像 (HorikawaGojo)



図 19: HorikawaGojo: サンプル画像

カメラは道路を跨ぐ歩道橋の上に設置し, 道路(堀川通り)の交通状況を撮影した(図 20). 撮影は, 直線道路を車両が画像下方から上方に進む様子を観測するように, 且つ, 互いの視野がほぼ重なるように行い(図 21), 照明変動, 道路状況, 機器設定を変化させて, 6 通りのデータセットを撮影した. 各データセットは 2700 フレーム × 2 視点計 5,400 枚の画像からなっている.

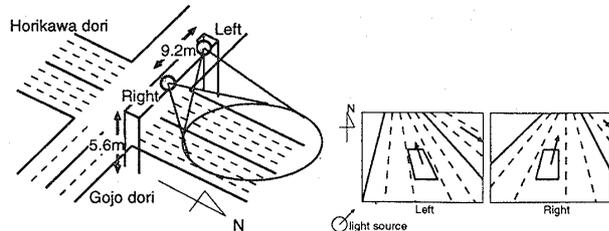


図 20: 物体, センサの配置図 図 21: シーンの概略図

- セット 1: 照明変動: 無(晴天)
道路状況: 通行量少→多, 遮蔽有
機器設定: ALC(OFF), EEI(OFF)

- セット 2: 照明変動: 無(晴天)
道路状況: 通行量少→多, 遮蔽無,
画面横から出てくる車両有
機器設定: ALC(OFF), EEI(OFF)
- セット 3: 照明変動: 有
道路状況: 通行量少→多, 遮蔽有,
画面横から出てくる車両有
機器設定: ALC(ON), EEI(OFF)
- セット 4: 照明変動: 無(曇天)
道路状況: 通行量少→多, 遮蔽有,
画面横から出てくる車両有
機器設定: ALC(ON), EEI(OFF)
- セット 5: 照明変動: 無(曇天)
道路状況: 通行量少→多, 遮蔽有,
路肩停止車両発進有
機器設定: ALC(ON), EEI(ON)
- セット 6: 照明変動: 無(晴天)
道路状況: 通行量少→多, 遮蔽無,
路肩停止車両発進有
機器設定: ALC(OFF), EEI(OFF)

4 まとめ

著者らは, 各種画像理解アルゴリズムの性能評価を客観的・科学的立場から行うための研究基盤整備として, 画像データベースの開発を行っている. 今回は, 最近注目されている多視点動画像を用いた動的シーン解析のための画像データセットの作成に主眼をおいて撮影を行った. 今後の撮影対象・方法としては,

- 従来のデータベースには比較的少ない屋外の多視点映像(道路シーン, 交通シーン)
- 赤外と可視光のマルチセンシングデータ(車両, 人物などを対象とする)
- 多くのグランドツルースを添付したデータ
- 撮影間隔を長くすることにより得られる長期間にわたる画像(例: 1 (frame/hour) で 1 年間撮影すると, 時間, 季節, 天候などの影響を観測できる)などを検討している.

本データベース開発は RWCP(新情報処理開発機構)知的資源 WG の活動として行ったもので, 同機構の許可を受ければデータの使用が可能である.

参考文献

- [1] 松山, 谷口 “Image Understanding Environment(IUE)プロジェクト”, 情処論誌, Vol.36, No.3, pp.203-211, 1995
- [2] 松山他 “IUE(画像理解用標準ソフトウェア)と Calibrated Image Database の現状”, 情処研報 CVIM 107-7, pp.49-56, 1997
- [3] 呂, 猪木 “ネットワーク化手話単語映像データベースの開発”, 信学技報 MVE 98-86, pp41-47, 1998
- [4] 和田, 浮田, 松山 “視点固定型パンチルトズームカメラとその応用”, 信学論, Vol.J81-D-II, No.6, pp.1182-1193, 1998