DCT を用いた合焦判定による物体検出法の提案

Detecting Objects by Using DCT Coefficients Analysis

正会員 三輪 道雄[†], フンフ クアンフイ ベト^{**}, 石井 雅博^{**}
小池 康晴^{**}, 正会員 佐藤 誠^{**}

Michio Miwa[†], Quang Huy Viet Huynh^{**}, Masahiro Ishii^{**},
Yasuharu Koike^{**} and Makoto Sato^{**}

Abstract The depth information of objects is important for retrieving and indexing image information. A method is described for finding the depth of objects by taking three pictures, each with a different focus. One picture is focused at the target position, one is focused near the target, and the last one is far from it. DCT coefficients are used to estimate how each picture is focused. By comparing the sum of the frequency coefficients of these pictures, we can determine whether an object is in the target position or not.

キーワード：合焦法，離散コサイン変換，画像処理，CCD カメラ，MPEG

1. ま え が き

監視カメラやロボット用のカメラにおいて，画像を撮影すると同時に撮影対象までの距離を知ることができれば，対象物の検出を容易に行うことができる．物体までの距離を知る方法のひとつとして，焦点を変えた複数の画像を撮影し，画像のぼけの大きさの差を利用する合焦法⁶⁾⁸⁾がある．この際，画像のぼけの大きさを測定するために，画素間の分散値を比較したり¹⁾，画像の持つ周波数成分を比較する．画像の持つ周波数成分を調べる方法としては，フーリエ変換を利用する方法⁷⁾¹⁴⁾があるが，離散コサイン変換(DCT)¹³⁾を利用した例はない．DCTを利用することができれば，画像の圧縮方法と共通に使用することができる．これは，動画像を記録するフォーマットとしてMPEG⁴⁾⁵⁾等が利用され，画像の圧縮にはDCTが用いられるからである．DCTにおいては，高速に行うことのできるアルゴリズムやハードウェアが利用可能であるため，距離計測を高速に行ったり，監視カメラなどのシステムを構築する際，システムの作成に必要な資源を少なくすることができる．

距離を測定する他の手法と比較した場合の合焦法の利点

1999年8月2日受付，1999年10月13日再受付，1999年11月5日採録

[†] 松下電器産業株式会社 マルチメディアシステム研究所
(〒140-8632 東京都品川区東品川4-5-15, TEL 03-5460-2744)

^{**} 東京工業大学 精密工学研究所
(〒226-8503 横浜市緑区長津田町4259, TEL 045-924-5050)

[†] Multimedia Systems Research Laboratory, Matsushita Electric Industries, Co., Ltd.
(4-5-15, Higashi-Shinagawa, Shinagawa-ku, Tokyo 140-8632, Japan)

^{**} P & I Laboratory, Tokyo Institute of Technology
(4259, Nagatsuta, Midori-ku, Yokohama 226-8503, Japan)

は，仕組みが単純になることである．すなわち，距離を測定する方法としては，アクティブな方法とパッシブな方法があり，アクティブな方法としては，レーザのスリット光を物体に向けて投射し，物体に投影されたパターンの形状から物体までの距離を測る光切断法²⁾がある．また，パッシブな方法としてはステレオ視³⁾があり，複数台のカメラの視差を利用して物体までの距離を測定する．アクティブな方法は正確であるが，カメラから物体までの距離が遠くなると，カメラ側から出すレーザ光等を強くする必要があり，環境に及ぼす影響が大きくなってしまふ．また，撮像系とは別に距離測定のためのしくみを用意する必要もある．一方，ステレオ視では，複数のカメラによって撮影した画像の中から同じ物体の点を対応付ける対応点問題を解く必要がある．合焦法はパッシブな方法のひとつであるが，焦点を変えて撮影した画像間の合焦の度合いを比較すればよく，カメラ側からレーザ光などを発生する必要はない．また，対応点問題を解く必要もない．

DCTを用いて合焦法により物体の位置を求める場合，画像に焦点ぼけが加わった際のぼけの大きさと，この画像にDCTを施して得られるDCT係数の関係を調べ，ぼけの大きさ変化を調べるためのDCT係数の使い方を決定する必要がある．また，合焦法は画像のぼけの大きさの変化によって物体までの距離を測定するものであるが，光学的な特性から，カメラから物体までの距離が遠くなるにつれて，焦点の合う範囲が広がってしまい，遠距離になるほど物体の位置を求めることが困難になってしまう．しかし，遠距離でも焦点の合う範囲を狭くすることができれば，特定距離にある物体の映像のみを取り出すことが単眼

でき、監視カメラなどに応用できる。

以下、本論文では、合焦法を用いて特定距離にある物体の検出を行う際の光学系について述べる。次に、筆者らの開発したスクリーン投影形撮像系^{9)~11)}による縮小光学系を用いれば、カメラから10 m程度離れた遠距離でも合焦法により物体位置の測定が可能であることを示す。さらに、画像のDCT係数が画像のぼけの大きさの判別に利用可能なことを数学的に示し、実験によってこれを確認する。また、従来用いられてきた離散フーリエ変換(DFT)との間に性能の差がないことを示す。なお、ぼけの広がり、幾何光学的には、有限円でモデル化することができる。しかし、レンズの特性が中央や周囲で異なるなどを考慮した場合、ガウス関数で表すことも可能である¹²⁾。本論文では、解析の目的に応じて、2章ではぼけの広がりを有限円でモデル化し、3章ではガウス関数を用いてモデル化している。

2. 一定距離範囲にある物体の検出

2.1 焦点を変えて撮影した3枚の画像による物体検出

焦点を変えた3枚の画像を撮影することにより、カメラから一定範囲内にある物体を検出する際の光学系を図1に示す。レンズから物体までの距離を u 、この物体の合焦位置にある投影面のレンズからの距離を v 、レンズの焦点距離を f とすると、レンズの公式(1)が成り立つ。

$$\frac{1}{f} = \frac{1}{u} + \frac{1}{v} \quad (1)$$

図1でレンズの左側を物体側、右側を投影面側と呼ぶことにする。投影面側に投影面l、投影面mおよび投影面nを配置して画像を撮影する。この時、レンズから各投影面までの距離をそれぞれ、 v_l 、 v_m 、 v_n とする。また、これらの投影面に合焦する物体までの距離をそれぞれ、 u_l 、 u_m 、 u_n とする。

u_l u_m の間に物体があり、レンズからこの物体までの距離を u_a とし、この物体の合焦位置のレンズからの距離を

図1 カメラモデル
A camera model.

v_a とする。 u_a にある物体の像は、投影面lおよび投影面m上ではぼけている。投影面におけるぼけの大きさの広がりをぼけ円と呼び、ぼけ円の投影面lおよび投影面m上の直径 r_{al} および r_{am} は、レンズの直径を D とすると、(2)および(3)式によって表すことができる。

$$r_{am} = \frac{|v_m - v_a|}{v_a} D \quad (2)$$

$$r_{al} = \frac{|v_l - v_a|}{v_a} D \quad (3)$$

$r_{am} = r_{al}$ となるとき、 u_a にある物体のぼけの大きさは投影面fおよび投影面l上で等しくなる。この時、 u_a は(4)式によって求めることができる。

$$u_a = \frac{f(v_m + v_l)}{v_m + v_l - 2f} \quad (4)$$

レンズから物体までの距離が u_a より大きくなると、投影面mにおけるぼけの大きさの方が投影面lにおけるぼけの大きさよりも大きくなる。一方、投影面mと投影面nに関して同じ議論が成り立ち、投影面n上での物体のぼけの大きさを v_n とし、投影面mおよび投影面n上でぼけの大きさが同じになる物体までの距離を v_b とすると、 v_b は(5)式によって求めることができる。

$$u_b = \frac{f(v_n + v_m)}{v_n + v_m - 2f} \quad (5)$$

物体までの距離が v_b より小さくなると、投影面mにおけるぼけの大きさの方が投影面nにおけるぼけの大きさよりも大きくなる。この結果、物体からレンズまでの距離 u が、 $u_a < u < u_b$ の時、投影面mにおけるぼけの大きさが投影面lや投影面nのぼけより小さくなり、物体がレンズから $u_a < u < u_b$ の距離にあることを検出することができる。

一方、 u_a 、 u_b が最初に与えられている場合、投影面を等間隔に配置するとすれば、(4)式および(5)式より、投影面の位置は(6)~(8)式となる。

$$v_f = -\frac{1}{2} \frac{f_{ub}}{u_b - f} + \frac{3}{2} \frac{f_{ua}}{u_a - f} \quad (6)$$

$$v_n = -\frac{1}{2} \frac{f_{ua}}{u_a - f} + \frac{3}{2} \frac{f_{ub}}{u_b - f} \quad (7)$$

$$v_m = \frac{v_f + v_n}{2} \quad (8)$$

なお、投影面上の像の大きさはレンズの焦点 F から投影面までの距離に比例して大きくなるので、ぼけの大きさを比較する際には、画像の拡大縮小が必要である。

2.2 スクリーン投影形撮像系

カメラから u の距離に焦点を合わせた時、 u の前後にある物体のぼけの大きさはぼけ半径 r として表現される。カメラの解像度を ϵ とした時

$$r < \epsilon \quad (9)$$

のとき、物体の像のぼけは観測されない。 u 前後の(9)式を満足する物体の移動範囲 Δ を物体深度と呼ぶ¹⁵⁾。物体深度は u の前方 Δ_1 と後方 Δ_2 があり、本稿では、レンズの焦点を f 、レンズの口径を D とした時、(10)式で表

すことができる。

$$\Delta = \frac{\Delta_1 + \Delta_2}{2} = \frac{\epsilon u f D (u - f)}{f^2 D^2 - \epsilon^2 (u - f)^2} \quad (10)$$

物体深度はレンズから物体までの距離が長くなればなるほど大きくなり、物体の位置検出が困難になる。物体深度を小さくするためには、レンズの口径 D を大きくするか、焦点距離 f を長くする必要がある。しかし、焦点距離を長くすると、投影面のサイズが同じ場合、視野角が狭くなる。投影面を大きくすることは CCD の面積を大きくすることに相当し、半導体の製造上、容易ではない。また、焦点距離を同じにしたままでレンズの口径を大きくすることは、レンズを構成するガラスなどの屈折率を大きくする必要があり、光学的なひずみが大きくなり、やはり困難である。

そこで、筆者らはスクリーン投影形撮像系⁹⁾を考案した。このカメラは主レンズによって得られる像をいったんスクリーンに投影し、この像を CCD カメラによって撮影するものである。スクリーン投影形撮像系の撮像系を図 2 に示す。この光学系はスクリーンを用いた縮小光学系であるが、スクリーンを用いることにより、大きな縮小率を容易に実現できる。また、レンズ群による縮小光学系では、対物レンズの像面に縮小用の光学レンズを置くため、視野角を広く保つためには、大きなレンズが必要となるが、スクリーンを用いればその必要がない。

スクリーン投影形撮像系の解像度 ϵ_s は、CCD の対角線の長さを d 、スクリーン上の CCD の撮影している範囲の対角線の長さを d_s とすれば (11) 式によって表される。また、スクリーン投影形撮像系の物体深度 Δ_s は、主レンズの口径を D_s 、主レンズの焦点距離を f_s とすれば (12) 式によって表される。

$$\epsilon_s = \frac{d_s}{d} \epsilon \quad (11)$$

$$\Delta_s = \frac{\epsilon_s u_s f_s D_s (u_s - f_s)}{f_s^2 D_s^2 - \epsilon_s^2 (u_s - f_s)^2} \quad (12)$$

この撮像系では、市販の CCD カメラを用いながら、主レンズとして口径が大きく焦点距離の長いレンズを用いることができ、投影面も大きくすることが可能ため、物体深度を浅くしかも視野角も広くすることが可能である。

図 3 は CCD カメラに直接レンズを取り付けた場合とス

図 2 スクリーン投影形撮像系
A screen projected optics.

クリーン投影形撮像系の比較である。CCD カメラに直接レンズを取り付けた場合の物体深度は (10) 式によって表される。この際想定しているのは、 $\epsilon = 0.018 \text{ mm}$ 、 $f = 30 \text{ mm}$ 、 $d = 12.7 \text{ mm}$ であり、レンズの性能を表す F 値 ($= \frac{\text{焦点距離}(f)}{\text{レンズの口径}(D)}$) を用いて表現している。

一方、スクリーン投影形撮像系で想定しているのは、CCD カメラは同じものを使用し、 $d_s = 126 \text{ mm}$ 、 $f_s = 299.2 \text{ mm}$ 、口径 $D_s = 53.4 \text{ mm}$ であり、 ϵ_s と ϵ の比は約 10 対 1 である。15 m 近辺で、物体深度が 2 m 程度になっており、 F 値として 0.6 以下のレンズを実現できたことになる。

今回の実験では用いなかったが、参考のために、スクリーン投影形撮像系を用いて 3 枚の画像を同時に撮影する撮像系を図 4 に示す。主レンズを通した像をビームスプリッタによって光路を分け投影面に投影する。投影された像を CCD カメラで撮影する。主レンズからスクリーンまでの距離は v_l 、 v_m 、 v_n に設定する。レンズから投影面までの距離を変えると投影面上の画像の大きさは変化するが、 C_1 、 C_2 、 C_3 を調整することにより、CCD カメラで撮影する像の大きさを同じにすることが可能である⁹⁾。

図 3 スクリーン投影形撮像系の性能例 (D の単位は mm)

Example of depth of field by the screen projected optics.

図 4 物体検出のための光学系
Optical system for objects detection.

3. DCT 係数を用いた合焦判定

画像に DCT を施し、その周波数係数を比較することにより、画像の相対的なぼけの大きさの比較を行う。最初に、焦点ぼけによる画像の周波数成分毎の減衰の様子を、1 次元関数を用いて検討する。次に、実際の 2 次元画像の合焦判定法について述べる。

3.1 DCT 係数のぼけによる影響

2 次元画像にぼけが加わった時のぼけの大きさと DCT との関係性を 1 次元波形について明らかにする。関数 $f(x)$ を $0 \leq x \leq 1$ で定義される積分可能な関数とし、コサイン変換形式で表す¹³⁾(13)式)。

$$f(x) = \sum_{n=0}^{\infty} a_n \cos(n\pi x) \quad (13)$$

ただし

$$a_n = \int_0^1 f(x) \cos(n\pi x) dx \quad (14)$$

関数 $f(x)$ を周期 2 で半周期毎に対称な $-\infty \leq x \leq \infty$ で定義される連続関数とし、 $f(x)$ を (15) 式および (16) 式を満足するように拡張する。

$$f(x+2L) = f(x) \quad (15)$$

$$f(x+2L-1) = f(1-x) \quad (16)$$

画像のぼけの大きさは、ガウス関数 ((17) 式) との畳込み積分で近似できる¹²⁾。

$$g(x, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \quad (17)$$

$f(x)$ と $g(x, \sigma)$ の畳込み積分によって得られる関数を $h(x, \sigma)$ とすれば、 $h(x, \sigma)$ は (18) 式になる。

$$\begin{aligned} h(x, \sigma) &= \int_{-\infty}^{\infty} g(x', \sigma) f(x-x') dx' \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x'^2}{2\sigma^2}} \sum_{n=0}^{\infty} a_n (\cos(n\pi x) \cos(n\pi x') - \sin(n\pi x) \sin(n\pi x')) dx' \\ &= \frac{1}{\sqrt{2\pi}\sigma} \sum_{n=0}^{\infty} a_n \left(\int_{-\infty}^{\infty} e^{-\frac{x'^2}{2\sigma^2}} \cos(n\pi x') dx' \right) \times \cos(n\pi x) \\ &= \sum_{n=0}^{\infty} a_n e^{-\frac{(n\pi)^2 \sigma^2}{2}} \cos(n\pi x) \end{aligned} \quad (18)$$

この結果は、 $f(x)$ のコサイン変換係数 a_n は $a_n e^{-\frac{(n\pi)^2 \sigma^2}{2}}$ となり、 n が大きいすなわち周波数が高いほど減衰が大きくなることを示している。

$f(x)$ を区間 $0 \leq x \leq 1$ で $x_i = \frac{2i+1}{16}$ (ただし、 $i=0, 1, \dots, 7$) で離散コサイン変換すれば、(18) 式で $n > 7$ は意味を持たない。この結果、(18) 式は (19) 式となる。

$$h(x_i, \sigma) = \sum_{n=0}^7 a_n e^{-\frac{(n\pi)^2 \sigma^2}{2}} \cos(n\pi x_i) \quad (19)$$

図 5 は (19) 式で表される a_n をすべて等しくし $\sigma=0$ 、 $\sigma=0.5$ および $\sigma=1.0$ にした結果を示している。また、図 6 はこれらの信号に DCT を施した後の、周波数成分を示している。周波数が高いほど、 σ を大きくした時の減衰が大

きいことがわかる。

以上は 1 次元の場合であるが、 8×8 画素で構成される 2 次元画像の場合は、(13) 式、(14) 式、(19) 式は (20) 式、(21) 式、(22) 式となる。

$$f(x, y) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} a_{nm} \cos(n\pi x) \cos(m\pi y) \quad (20)$$

ただし

$$a_{nm} = \int_0^1 \int_0^1 f(x, y) \cos(n\pi x) \cos(m\pi y) dx dy \quad (21)$$

$$\begin{aligned} h(x, y, \sigma) &= \sum_{n=0}^7 \sum_{m=0}^7 a_{nm} e^{-\frac{(n^2+m^2)\pi^2 \sigma^2}{2}} \\ &\quad \times \cos(n\pi x) \cos(m\pi y) \end{aligned} \quad (22)$$

3.2 合焦判定アルゴリズム

3.1 節に示すように、周波数成分が高いほどぼけによる減衰が大きい。しかし、撮影された画像の持つ周波数成分は対象に依存するので、合焦判定に適した帯域は実験で求めることとし、本節では対象物を検出するためのアルゴリズムと合焦判定に用いる帯域の設定方法について述べる。

8×8 画素で構成される焦点を変えて撮影した 3 枚の 2 次元画像を $f^{(i)}(x, y)$ ($i=1, 2, 3$) とし、 $f^{(i)}(x, y)$ に DCT を施して得られる係数を $a^{(i)}(n, m)$ とする。 i が大きくなるに従って焦点を合わせるカメラから物体までの焦点距離が遠いとする。そして、(23) 式で与えられる $v^{(i)}$ を計算す

図 5 ぼけの影響による波形の変化
Test signal and blurred signal.

図 6 DCT 係数のぼけによる減衰
Coefficients of DCT of test signal.

図 7 DCT 係数配列
Array of DCT coefficient.

る.

$$v^{(i)} = \sum_{n=0}^7 \sum_{m=0}^7 o(n, m) (a^{(i)}(n, m))^2 \quad (23)$$

$$o(n, m) = \begin{cases} o(n, m) = 1 & \text{if } p_l \leq m + n \leq p_h \\ o(n, m) = 0 & \text{other} \end{cases}$$

ここで DCT 係数のうち合焦判定に用いる p_l と p_h によって指定されるバンド幅の指定方法について述べる. 図 7 は 64 個の DCT 係数の配列を表しており, 左上が直流成分, 右下が高周波成分である. この時, 右上から左下に向かう直線上の周波数成分 $a(n, m)$ の n, m の和は一定であり, これを p とする ($p = n + m$). p_l は直流成分に近い直線上, p_h は高周波成分に近い直線に対応している. p_l と p_h は実験によって決定する.

最後に, 3 つの $v^{(i)}$ に関して (24) 式が成り立った時, $f^{(i)}(x, y)$ として撮影された対象物が目的距離にあると判断する.

$$v^{(2)} > v^{(1)} \quad \text{かつ} \quad v^{(3)} \quad (24)$$

4. 実験

4.1 スクリーン投影形画像観測系を用いた画像の撮影

実験では, CCD カメラを 1 台とし, 主レンズとスクリーンの距離を変えて 3 枚の画像を取得することとし (図 2 参照), 表 1 に示す機材を用いてスクリーン投影形撮像系を作成した⁹⁾. スクリーン上の CCD カメラの撮影している範囲は $d_s = 126 \text{ mm}$, F 値が 0.6 以下のレンズと同じになる. また, この撮像系では視野角 (ω) 23.8° となる.

目的距離 (10 m) に物体を配置し, 遠距離 (16 m), 目的距離 (10 m), 近距離 (7 m) に焦点を合わせて 3 枚の画像を撮影する. この結果, 8.2 m から 12 m の範囲 (目標距離範囲) にある物体が検出できることになる. この範囲は, 時速 4 km で歩いている人間を手の範囲を含めて 1 サイクル/秒で捕捉できる範囲をもとに設定した. 目標物が撮影されている範囲を 8×8 画素の小領域に分割し (合計 4096 個), 周波数成分を比較する.

4.2 実験および考察

合焦判定に適したバンド幅を決定するために, 試作したカメラを用いて図 8 に示す 2 つの目標物を目的距離において画像を撮影した. それぞれ 105 個, 96 個の小領域で構成されている.

図 9 は目標物 (a) について p_l と p_h を変えながら小領域

表 1 試作したスクリーン投影形画像観測系の仕様
Specification of experimental option.

レンズ (Nikon W, 300 mm, F 5.6)	
F ナンバー	5.6
焦点距離	299.2 mm
スクリーン (タチカワ フレネルレンズ)	
サイズ	202.4 mm $\times$ 253 mm $\times$ 2.1 mm
CCD カメラ (Panasonic GP-US 522 HP)	
焦点距離	16 mm
受光面サイズ	12.7 mm
画素数	512 $\times$ 512
解像度	256 階調 (グレースケール)

図 8 目標物
Target objects.

図 9 バンド幅の変化による正解率の変化
Rate of correct detection according to the DCT coefficient usage.

図 10 領域抽出結果-1(左：手前，中：正解，右：遠く)
Result of evaluation 1 (left: judged near, middle: correct, right: judged far).

図 11 領域抽出結果-2(左：手前，中：正解，右：遠く)
Result of evaluation 2 (left: judged near, middle: correct, right: judged far).

の判定を行い，小領域が目標距離範囲内と判別された時の正解率を示している． $p_l=3$ ， $p_h=4$ のとき最も正解率が高く 80%以上である．直流成分に近い係数を用いない方が良い結果が得られているのは，ぼけによって周囲のブロックの影響が及んだ場合，直流成分に近い係数に影響があるためと考えられる．また，高周波の係数を除いた方が良い結果が得られるのは，今回撮影した対象物および撮像系では，得られた画像に高周波が含まれておらず，高周波の係数のみを使用すると，ノイズの影響などで判別に誤りが発生するためと考えられる．

同じ評価を目標物(b)についても行い，これらを総合して $p_l=3$ ， $p_h=6$ を採用することにした．図 10，図 11 は $p_l=3$ ， $p_h=6$ のとき，目標物の小領域が目標距離範囲内，手前，遠くのどこに判定されたかを示している．どちらの画像でも，模様の顕著でないところで判別エラーが発生している．

次に廊下に物体を配置し評価した．図 12 は廊下の風景である．枠の中は目標距離範囲を示している．目標物を含む風景の中から目標距離範囲にあるものを検出したものが図 13 である．図 14 は判定結果をブロック単位に色分けしたものである．黒いブロックが目標距離範囲にあると判定されたブロックで，白いブロックが目標距離範囲外にあると判定されたブロックである．

図 13 では，多数の目標距離範囲外のブロックが目標距離範囲と判別されている．これは，対象物がテクスチャを

図 12 廊下の風景
A scene of corridor.

図 13 廊下のシーンからの検出結果(閾値なし)
Result of evaluation (no threshold).

多く持たないとき，DCT 係数の値が小さくなって誤った検出が行われていると考えられる．

そこで， V に閾値を設け，各ブロックの 3 つの V のどれもがこの閾値を超えない時は判定から除外する．図 15，図 16 は，すべてのブロックの V の中の最大値の 50% を閾値として，判定を行ったものである．図 15 では，目標距離内にある人の画像が検出されていることがわかる．図 16 で黒ブロック白ブロック以外は閾値より小さく判定されなかった部分である．判別対象となったブロックは全体の 19% であり，目標距離内の正解率は 73%，目標距離外の正解率は 76% であった．廊下の壁などで，同じ小領域の中に，距離の異なったものが含まれている場合，誤認識が多くなっている．

DFT との性能の比較をするために，目標物(a)および目標物(b)について，バンド幅を変えながら DFT 係数の

図 14 廊下のシーンからの検出結果パターン(閾値なし)
Result of evaluation pattern (no threshold).

図 16 廊下のシーンからの検出結果パターン(閾値あり)
Result of evaluation pattern (using threshold).

図 15 廊下のシーンからの検出結果(閾値あり)
Result of evaluation (using threshold).

実部と虚部の 2 乗和を比較して正解率を求めた。目標物 (a) の場合、最大の正解率は DCT の場合 81%, DFT の場合 80% であった。目標物 (b) では、DCT, DFT 共に 77% で同じであった。

5. む す び

DCT 係数を用いたぼけの大きさの測定を使用して、合焦法により物体位置の検出が可能であることを示した。また、DFT と比較して同程度の性能が得られることが確かめられた。DCT によってぼけの大きさを測定することができれば、画像の圧縮と同時に物体までの距離推定を行うことが可能である。ぼけの大きさの測定に利用する DCT の係数を制限することにより、対象物が平面の場合は約 80% の精度で対象物が目的距離範囲内にあるか否かを識別できた。廊下を対象とした実験では、周波数成分を多く含

まないブロックを判定から除外するために、閾値を設けることにより 73% のブロックが目標距離内にあると識別できた。また、スクリーン投影形撮像系を用いることにより、従来の光学系では困難であった 10 m 以上の距離における物体検出が可能となった。判定領域の形状は、今回用いた斜め直線以外に、円弧形が考えられる。図 7 で、各係数の左上の直流成分の係数からの距離によってバンド幅を決め、正解率を求めた。その結果、目標物 (a) では 81%, 目標物 (b) では 80% の正解率が得られ、今回用いた斜め直線によって判定領域を設定する場合とほぼ同じ結果が得られた。最適な判定領域の形については、対象となる画像に依存すると考えられ、さらに厳密な検討が必要である。

課題としては、今回の実験では、ぼけの大きさの評価に用いる DCT 係数のバンド幅は撮影対象物 2 つの目標物についての実験で決定しているが、適切なバンド幅は撮影対象物によって異なると考えられる。そこで、撮影対象物によって評価に用いるバンド幅を可変にする方法を検討する必要がある。また、ブロック内に距離の異なったものが撮影されている場合、判定の誤りが発生するため、隣接ブロックとの関連を考慮した判定が必要である。今後はリアルタイムの物体検出をするために、3 枚の画像を同時に撮影する撮像系の実現をしていきたい。

[文 献]

- 1) 工藤朋之, 三池秀敏: “カメラフォーカスを連続変化させた動画像を用いた奥行き分布検出”, 情処, CV 98-2, pp. 7-12 (Jan. 1996)
- 2) 井口, 佐藤: “三次元画像計測”, 昭光堂 (1990)
- 3) 吾妻健夫, 魚森謙也, 森村淳: “ステレオ画像の中間視点画像生成のためのエッジ情報を用いた視差推定”, 映像学誌, 52, 3, pp. 322-330 (Mar. 1998)
- 4) ISO, “Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbits/s Parts Video”, ISO/IEC 11172-2
- 5) ISO, “Coding of moving pictures and associated audio Part 2 Video”, ISO/IEC 13818-2 (Committee Draft)
- 6) A. Pentland: “A New Sense for Depth of Field”, IEEE Transaction

- on Pattern Analysis and Machine Intelligence, **PAMI-9**, 4 (July 1987)
- 7) A. Pentland, T. Darrell: "A Simple, Real-time Range Camera", Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, California (June 1989)
 - 8) Masahiro Watanabe, Shree K. Nayer: "Minimal Operator Set for Passive Depth from Defocus", Proc. of IEEE Conf. Computer Vision and Pattern Recognition, pp. 431-438 (1996)
 - 9) 大原知之, 三輪道雄, 石井雅博, 小池康晴, 佐藤誠: "合焦法距離計測のためのスクリーン投影型画像観測系の提案", 信学論 D-II, **J82**, 10, pp. 1750-1757 (Oct. 1999)
 - 10) 三輪道雄, 大原知之, 石井雅博, 小池康晴, 佐藤誠: "焦点ボケを用いた遠距離物体領域の抽出法", 第 5 回画像センシングシンポジウム論文集, pp. 221-226 (June 1999)
 - 11) 三輪道雄, 大原知之, 石井雅博, 小池康晴, 佐藤誠: "焦点ボケを用いた遠距離物体の推定", 3 次元画像コンファレンス '99, pp. 305-307 (June 1999)
 - 12) Murali Subbarao, Gopal Surya: "Depth from Defocus: A Spatial Domain Approach", International Journal of Computer Vision, **13**, 3, pp. 271-294 (1994)
 - 13) K. R. Rao, P. Yip 著, 安田, 藤原訳: "画像符号化技術", オーム社 (1992)
 - 14) 日浦, 松山: "構造化瞳をもつ多重フォーカス距離センサ", 信学論 D-II, **J82**, 11, pp. 1912-1920 (Nov. 1999)
 - 15) 三宅和夫: "幾何光学", 共立出版 (1979)

三輪 道雄 (みわ みちお) 1979 年, 名古屋大学理学部物理学学科卒業, 同年, システムサービス(株)入社. 1984 年, 松下電器産業(株)入社, 現在に至る. 以来, 画像処理ヒューマンインタフェースの研究に従事. 1999 年 9 月東京工業大学総合理工学研究科知能システム科学博士課程(社会人入学)満期退学. 正会員.

Huynh Quang Huy Viet (フュンフ クアン フイ ベト) 1995 年, ホーチミン工科大学卒業. 1998 年, 東京工業大学研究生, 1999 年より東京工業大学修士課程在学中.

石井 雅博 (いし い まさひろ) 1990 年, 秋田大学電子工学科卒業. 1995 年, 東京工業大学大学院博士課程修了. 同年, 同大精密工学研究所助手, 現在に至る. ヒューマンインタフェース, 心理物理学の研究に従事. 工学博士.

小池 康晴 (こ いけ やすはる) 1987 年, 東京工業大学工学部情報工学科卒業. 1989 年, 東京工業大学総合理工学研究科修士課程修了. 同年, トヨタ自動車(株)入社. 1991 年, ATR 視聴覚機構研究所出向. 1992 年, ATR 人間情報通信研究所. 1998 年, 東京工業大学精密工学研究所助教授. 運動制御の神経回路機構の研究に従事. 工学博士.

佐藤 誠 (さとう まこと) 1973 年, 東京工業大学工学部電子物理工学科卒業. 1978 年, 同大学院博士課程修了. 同年, 同大学工学部助手. 現在, 同大学精密工学研究所教授. パターン認識, 画像処理の研究に従事. 工学博士. 正会員.