

UDC 681.3: 685.5
621.9-523.8: 621.397.13

論文
51-C7

移動物体の実時間視覚認識方式とその応用

日立製作所
日立製作所
日立製作所

烏野 武
江尻 正員
徳永 趟

1. まえがき

生産工程の自動化においては、対象となる物体の形状や位置を認識できる視覚装置の開発が一つの重要な課題である⁽¹⁾⁽²⁾。

物体の認識の研究は人工知能の分野におけるシーンアナリシスの一環として数多くの研究が行なわれてきており、その中のいくつかは人工知能ロボットのプロトタイプへと開発が進んで来ている^{(3)~(5)}。しかしながら、実用的なシステムへ発展させるにはいくつかの理由によりいまだに困難を極めている。

その理由は、第1に対象の認識に要する処理時間が膨大になること、第2に認識アルゴリズムが複雑になって一般に大規模な処理装置を要すること、第3に対象の映像パターンを取得するための撮像装置が雑音、ドリフト、ひずみなどの不安定性を含み十分なものでないこと、である。特に移動する対象を実時間で認識するには、撮像装置の残像現象による像ぶれが問題となり、これが認識をますます困難にしている。

本研究ではこのような移動物体を認識の対象とし、これに対して計算機を用いない単純な方法で高速かつ安定に認識する方式を考察した。この認識装置は物体の像が映像中を1次元的に移動する場合を対象とし、物体像の形状を判定するとともに、設定された場所を通過するタイミングを出力するものである⁽⁶⁾。この方式では、映像中に設けた複数個の窓わく内の情報の組合せによって物体の形状を判定するが、ここで用いる

A Method of Real-time Recognition of Moving Objects and Its Application. By Takeshi Uno, Member, Masakazu Ejiri, Member (Central Research Laboratory, Hitachi Ltd.) & Takeshi Tokunaga (Taga Works, Hitachi Ltd.).

鳥野 武：正員、日立製作所中央研究所
江尻正員：正員、日立製作所中央研究所
徳永 趟：日立製作所多賀工場

演算は窓わく内のある状態をもった絵素の計数、すなわち積分形の演算を基本としているため、雑音や形状の変化の影響を受けにくい方法となっている。

本論文では、この認識方式の原理と、これを視覚装置として実現し、コンクリート成形わくのボルト締継作業の自動化に適用した結果について述べる。

2. 移動物体の映像とその性質

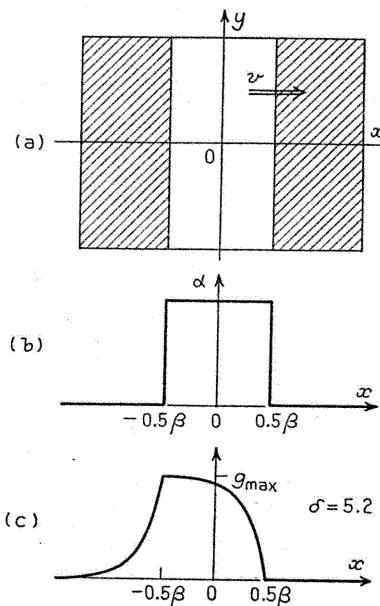
まずビジコン・カメラのような残像現象のある撮像装置で移動する物体を撮像した場合の映像信号について考察する。簡単にするために、撮像装置のレンズ系を介した後の物体映像と、これより得られる映像信号との関係について考える。

いま、映像面の座標 (x, y) 、時刻 t における映像の明るさを $f(x, y, t)$ とすれば、残像現象のある撮像装置での映像信号 $g(x, y, t)$ は近似的に(1)式で表わすことができる。

$$g(x, y, t) = k \int_{-\infty}^t f(x, y, \tau) e^{-\lambda(t-\tau)} d\tau \quad (1)$$

ここに、 k は比例定数であり、 λ は撮像装置の残像特性に関わる映像信号の減衰率である。すなわち、(1)式は任意の点における時刻 t の瞬間の明るさの重み1に対し、それ以前の映像の明るさは時間とともに指数関数的に減少しつつ映像信号に残像として影響することを表わしている。

そこで、第1図(a)のように、黒色背景の前を白色物体の映像が x 方向に定速度 v で移動する場合を考える。(b)図のような映像の黒色部分の明るさを0、白色部分の明るさを α 、白色部分の x 方向の幅を β とし、 $t=0$ の瞬間に白色部分の中心が $x=0$ の点を通るように原点をとると、映像の明るさは(2)式で表わせる。



第1図 移動白色物体像とその映像信号波形

Fig. 1. Image and video signal of a white moving object which projects a striped pattern.

$$f(x, y, t) = \begin{cases} \alpha, & \cdots -\beta/2 + vt \leq x \leq \beta/2 + vt \text{ の} \\ & \text{とき} \\ 0, & \cdots \text{その他のとき} \end{cases} \quad (2)$$

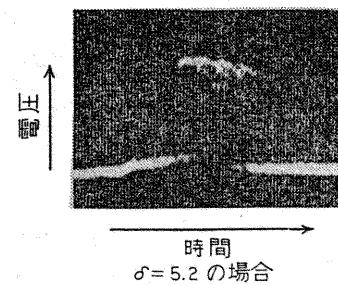
t と $(x - \beta/2)/v, (x + \beta/2)/v$ の大小関係を考慮しつつ、(2)式を(1)式に代入し積分を実行し、かつ $t=0$ として(b)図に対応する時点における映像信号を求めると(3)式を得る。

$$g(x, y, 0) = \begin{cases} \frac{k\alpha}{\lambda} \{ e^{\delta(x/\beta+1/2)} - e^{\delta(x/\beta-1/2)} \}, & \cdots x/\beta < -1/2 \text{ のとき} \\ \frac{k\alpha}{\lambda} \{ 1 - e^{\delta(x/\beta-1/2)} \}, & \cdots \\ 0, & \cdots 1/2 \leq x/\beta \leq 1/2 \text{ のとき} \end{cases} \quad (3)$$

ただし、 $\delta \equiv \lambda\beta/v$

これは(c)図のような波形の映像信号を意味しており、この波形は実際にビデオ・カメラで水平方向に移動する白色物体を撮像したときに観測される波形に良く一致する。第2図は特定の一水平走査線の映像信号を数フィールドにわたって重畠撮影した写真である。映像の該当部分が 1/60s ごとに走査されることから、第1図(c)のような波形が間欠的に移動していく様子を見ることができる。

さて、(3)式から映像信号は $x/\beta = -1/2$ のときに



第2図 移動白色物体の映像信号波形

Fig. 2. Actual video signals for a white moving object.

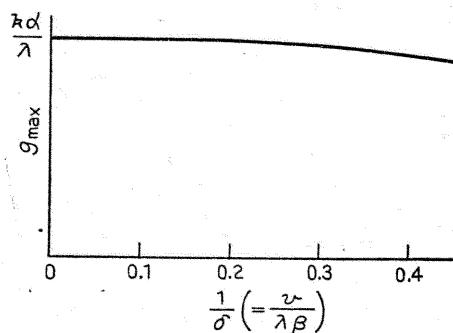
第3図 移動白色物体の映像信号の最大値 g_{\max} と δ の関係

Fig. 3. Relationships between the maximum image value g_{\max} and δ .

最大となり、最大値 g_{\max} は(4)式のように導かれる。

$$g_{\max} = \frac{k\alpha}{\lambda} (1 - e^{-\delta}) \quad (4)$$

一般に物体が移動している場合の g_{\max} は静止している場合のそれよりも小さいが、 δ が大であれば、すなわち移動速度 v に比べて撮像装置の減衰率 λ が大きいか、あるいは物体像の横幅 β が大きければ、第3図に示すように g_{\max} はほとんど v の影響を受けない。後述の応用例においては δ は 3.3 ないし 11.2 程度であるので、最も影響の著しい $\delta = 3.3$ の場合でも、 g_{\max} は物体が静止している場合のそれの約 96% となるに過ぎない。

また、(3)式で表わされる映像信号が、

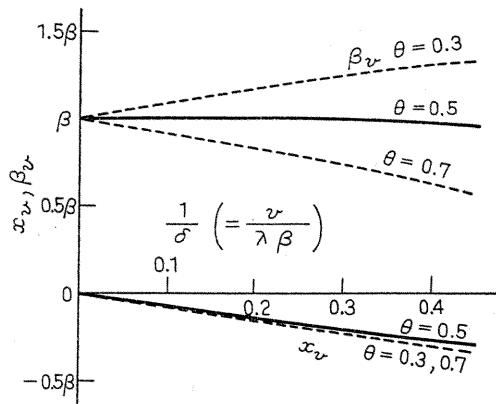
$$g \geq \theta \frac{k\alpha}{\lambda}, \quad \text{ただし, } 0 < \theta \leq 1 - e^{-\delta} \quad (5)$$

を満足する部分の横幅 β_v 、この部分の中心位置のずれ x_v はそれぞれ(6)、(7)式で与えられる。

$$\beta_v = \beta + \frac{\beta}{\delta} \log_e \frac{(1 - e^{-\delta})(1 - \theta)}{\theta} \quad (6)$$

$$x_v = -\frac{\beta}{2\delta} \log_e \frac{1 - e^{-\delta}}{\theta(1 - \theta)} \quad (7)$$

すなわち、物体が静止している場合の映像信号の最大



第4図 移動物体像(2値)の横幅 β_v 、中心位置のずれ x_v と δ の関係

Fig. 4. Relationships between the image width β_v , image center position x_v and δ .

値の θ 倍をしきい値として2値の映像信号を作るとき、第4図に示すように、 θ を 0.5 とし δ が大であれば移動中の白色物体像の横幅 β_v はほとんど静止時の横幅 β に等しい。また物体像中心の位置ずれ x_v は δ が大であれば v/λ に比例する。この比例係数は θ が 0.5 のとき最も小さい。しかし、 θ を 0.5 を選んだとしても δ が十分大でなければ、位置ずれ x_v を無視することはできない。

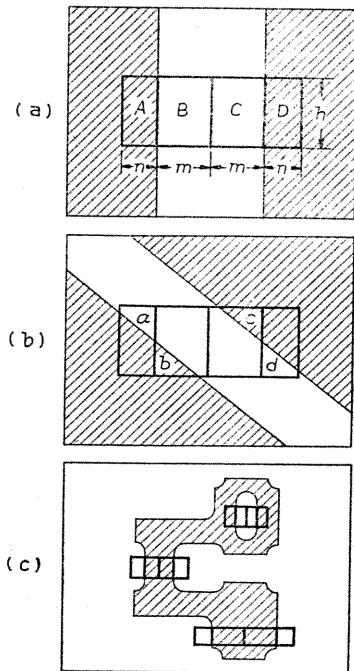
3. 認識方式

<3・1> 基本構成 以上の考察から、物体の映像はその物体が移動していようと適切なしきい値を選んで2値化することにより、ほぼ一定の幅を保たせ得ることがわかった。従って、検出すべき物体をその背景から明確に区分できるように照明などの外的条件を整えることにはれば、認識が容易になるばかりでなく2値パターンを扱うだけで良いから処理が単純化できるなど、実用化するうえで効果的である。

さて、上述のようにして2値化した帯状物体像に対し、第5図(a)のように四つの小領域 A, B, C, D から構成される窓わく領域を考える。窓わくの小領域 B や C の横幅は等しく m とし $2m=\beta$ に選ぶ。また小領域 A や D の横幅は等しく n とし、 n は任意に選び、窓わくの高さ h も任意に選ぶ。

帯状物体像がこの窓わくの中央にきた瞬間を考えると、小領域 B と C は物体で覆われ A と D は背景で覆われることになる。一般に(b)図に示すような何らかの物体が入ってきたとき、 A 中の背景でない面積 a 、 B 中の物体でない面積 b 、 C 中の物体でない面積 c 、 D 中の背景でない面積 d により、

$$p=a+b+c+d \dots \dots \dots (8)$$



第5図 移動物体像の形状位置認識用窓わく
Fig. 5. Windows for recognizing the position and shape of a moving object.

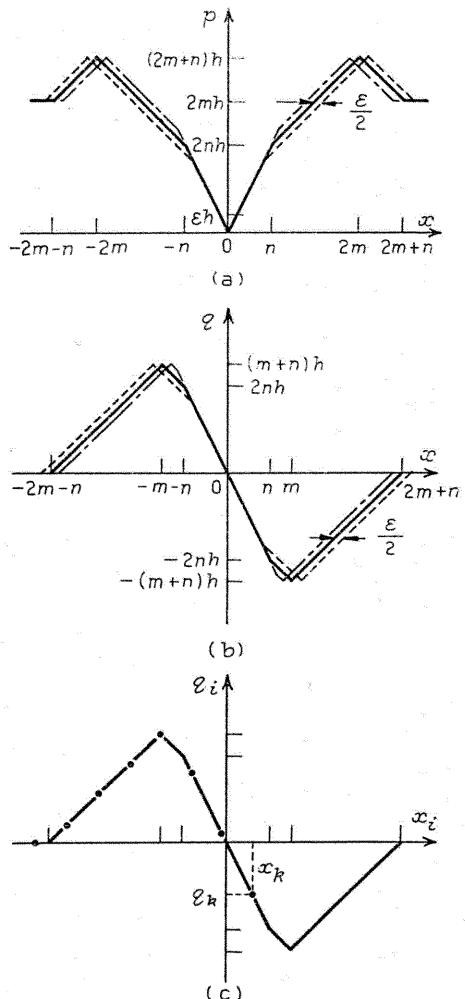
$$q=a-b+c-d \dots \dots \dots (9)$$

を計算する。

このとき、数値 p は(a)図の窓わく領域内のパターンと対象物体の窓わく領域内のパターンとのずれの程度を与える。また数値 q は、 A および B 中の物体の面積 $a+(mh-b)$ と C および D 中の物体の面積 $(mh-c)+d$ との差に等しいから、窓わく領域内の対象物体の平均的な中心と窓わくの中心との位置ずれの程度を与える。

<3・2> 形状の認識 以上の原理をもとに物体の形状を認識するには次のようにすればよい。まず、第5図(c)に示すように物体の特徴的な部分に着目し、それぞれの横幅に対応して複数個の窓わくを設定し、(8)式によってそれぞれの窓わくについて p を計算し、これらがそれぞれ微小であったときに対象物体の形状が所定のものと一致すると判定する。もちろん、対象物体は移動するものと考えているから、上記の計算と判定は刻々変化する映像に対して繰返し実施することになる。

いま、物体像が x 方向に移動する場合を考え、ある窓わくについての(8)式による p と、物体像の中心位置 x の関係を求めると、第6図(a)の実線のようになる。もし物体像の横幅が $\beta+\epsilon$, $\beta-\epsilon$ ($\epsilon > 0$ とする) ならば、それぞれ同図の破線、鎖線のようになる。少なくとも $x=0$ の近傍では、



第 6 図 窓わく内の物体像から求めた p , q と物体像の中心位置 x の関係

Fig. 6. Values p and q in relation to the object center position x .

$$p(x) = \begin{cases} -2hx, & -n + \frac{\epsilon}{2} \leq x < -\frac{\epsilon}{2} \\ \epsilon h, & -\frac{\epsilon}{2} \leq x \leq \frac{\epsilon}{2} \\ 2hx, & \frac{\epsilon}{2} < x \leq n - \frac{\epsilon}{2} \end{cases} \quad \dots \dots \dots (10)$$

ただし, $0 \leq \epsilon < n$

が成り立つので, 形状の許容誤差 ϵ , を設定し, 未知の物体像の移動に従って連続的に ϵ を観測するとき, $p \leq \epsilon, h$ となるならば, この未知の物体像の平均的な横幅と標準物体の横幅との差は $\pm \epsilon$, 以内であると言える。

従って複数個の窓わくに対し, それぞれの ϵ が同時に設定された値以下になったことを検出して, 移動中の物体が所定のものであるという判定を行なえば良

い。

〈3・3〉 位置の認識 物体像が映像面の所定位置に設定された窓わくの中心を通過する時点を検出するには, 上述の ϵ が最初にしきい値 ϵ, h 以下になる時点として検出することもできるが, (9)式による q を用いたほうがより正確に検出できる。 q は移動する物体像の中心位置 x の関数として第6図(b)のようになる。ここに実線, 破線, 鎖線は同図(a)に準じて表わす。この場合 $q(x)$ は少なくとも $x=0$ の近傍で,

$$q(x) = -2hx, \dots \dots \dots$$

$$-n + \frac{\epsilon}{2} \leq x \leq n - \frac{\epsilon}{2} \text{ のとき } \dots \dots \dots (11)$$

ただし, $0 \leq \epsilon \leq 2n$

なる1次の関数となるので, 物体像が x 方向に移動する場合 q が正から負に転じる時点をもって位置を検出すれば, ϵ の影響を避けることができる。

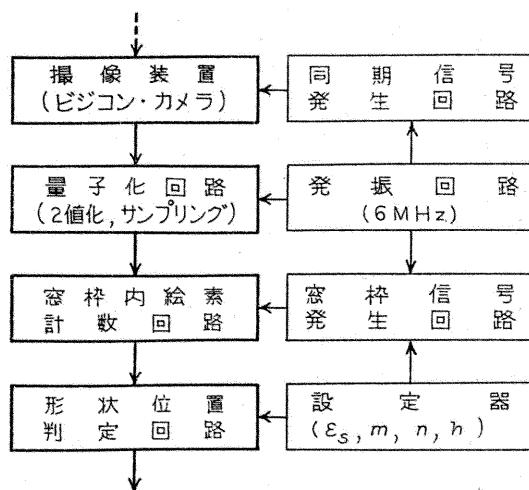
しかしながら, 一般に p , q は1映像につき1回しか演算できない間欠的な観測値である。そこで, i 番目の映像における x , p , q をそれぞれ x_i , p_i , q_i とし, 第6図(c)に示すように k 番目の映像において $q_{k-1} > 0$ であり, かつ $q_k \leq 0$ が成立した場合, (10)式から導かれる $p_k \leq \epsilon, h$ を満たせば, 未知の物体像は設定された物体像の条件に合致すると判定し, また未知の物体像の中心位置 x_k は, (11)式から導かれる $x_k = -q_k/2h$ により求めるようとする。ただし, ここでは物体像の横幅の誤差の許容値 ϵ , は n より小さく設定されなければならず, また1映像周期中の物体像の移動距離は $\epsilon/2$ 以下でなければならない。

なお, 第5図(c)に示すように複数個の窓わくを用いる場合でも, 位置の判定はそれらの中の代表的な一つの窓わくについて行なえば十分である。もし代表的な二つの窓わくについて位置の判定を行なえば, 窓わくの位置関係とそれぞれについて算出される x_k から物体像のある程度の傾きも検出できることになる。

4. 視覚装置の構成

以上の認識方式を第7図に構成線図で示すような電子回路で構成した。この装置は, ビジコン・カメラからの映像信号を2値化するとともに, サンプリングにより絵素化する量子化回路と, 設定器からの情報で窓わく信号を発生する回路と, この窓わく内部の特定の状態をもつ絵素を計数する回路と, この計数値から求めた ϵ や q から形状と位置を判定し出力する回路からなる。

ここで, 映像信号のサンプリング周期は 167 ns (6



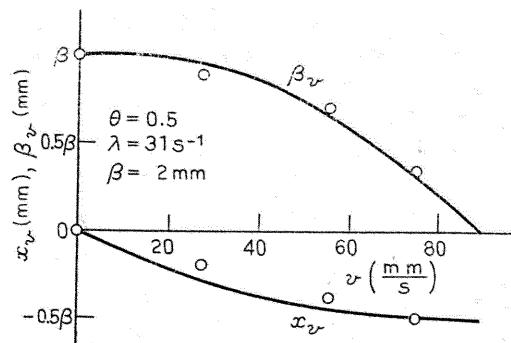
第7図 視覚装置の構成線図

Fig. 7. Block diagram of an image processor.

MHz) としており、382 回のサンプリングごとに水平同期信号を、262 回の水平同期信号ごとに垂直同期信号を発生し、ビジコン・カメラを駆動している。これにより 382×262 個の絵素を得るが、このうち帰線期間を除いた 320×240 絵素が有効な映像情報をもつ。この方式ではインターストロブ走査を行わないために、映像入力周期は 16.7 ms となり、形状位置認識もこの周期で動作することになる。

窓わく内絵素計数回路は窓わく信号発生回路からの信号に従い、窓わく内の映像を走査している期間のみ窓わく内部の特定の状態をもつ絵素の数を計算し a, b, c, d を求めるものである。また形状位置判定回路は i 番目の映像のこれらの数値から p_i と q_i を計算し、 $q_{i-1} > 0, q_i \leq 0, p_i \leq \epsilon, h$ のすべてを満たすかどうかを判定する。 k 番目の映像でこの 3 条件が満たされたとき、前述の原理で q_k から x_k を求めるのであるが、実際にはこの x_k をそのまま出力するのではなく、映像周期の $1/4$ にあたる 4.2 ms の周期パルスをあらかじめ発生しておき、 x_k の値を加味して送出パルスの選択を行なう方法を探っている。具体的には、 k 番目の映像入力時点から 1 映像周期分遅れて発生するタイミング・パルスを基準とし、この基準タイミング・パルスよりほぼ x_k/v だけ早いタイミングに発生するパルスを選択して送出している。これにより物体像の中心が設定した窓わく中心を通過してからほぼ 1 映像周期遅れてパルスが送出されることになるが、このことは物体像の移動が等速である限り、通過検出位置を窓わく中心から 1 映像周期間の物体像移動距離だけ x 方向にずらしたことと等価であり、実用上の問題はまったくない。

なお本視覚装置では、一つの標準物体像について 3

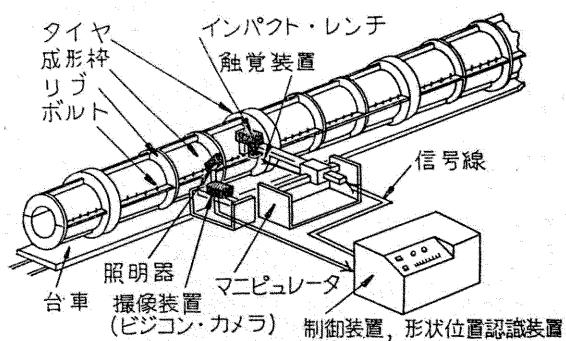
第8図 ビジコン面上における移動物体像 (2値) の横幅 β_v 、中心位置のずれ x_v と δ の関係Fig. 8. Object width β_v and its center position x_v on vidicon surface.

個の窓わくを設定できるようにし、形状位置判定回路は各窓わくについての p, q 値を一時記憶し、すべての窓わくについての処理が完了した時点で総合的な判定を行なっている。

また本視覚装置では、窓わく信号発生回路から形状位置判定回路までを 2 系統内蔵し、2 種類の物体を識別できるようにしている。この装置における物体像のビジコン面上での移動速度 v に対する横幅 β_v と真の中心位置からのずれ x_v との測定値を第 8 図に示す。この図は第 4 図と等価なものであり、 $1/\delta$ が $0 \sim 1.45$ の範囲にあたる。また測定速度の上限は後述の応用例における移動速度の約 8.5 倍にあたるが、図に実線で示す (6), (7) 式に基づく理論値と良く一致している。

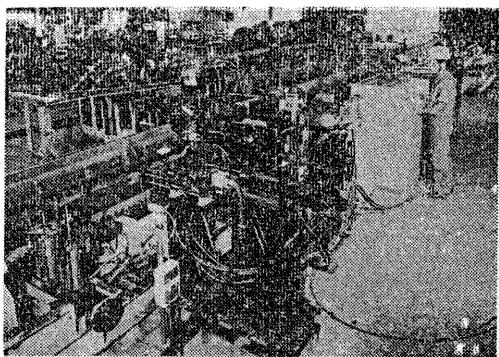
5. 視覚装置の応用

以上述べた視覚装置をコンクリート・パイルやポール製造の前工程や後工程である成形わくのボルト締緩作業に適用した。その全体構成を模式的に第 9 図に示す。また全体の装置写真を第 10 図に示す。



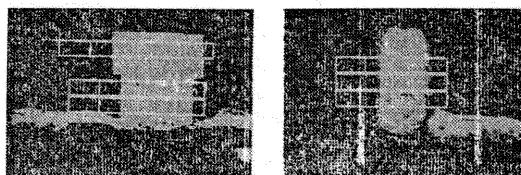
第9図 コンクリート・ポール成形わく用自動ボルト締緩装置の構成

Fig. 9. Configuration of the bolting robot for concrete manufacture.



第 10 図 コンクリート・ポール成形わく用
自動ボルト締緩装置

Fig. 10. General view of the bolting robot.



第 11 図 コンクリート・ポール成形わくの
ボルト、リブ像とその形状位置
認識用窓わく

Fig. 11. Object images and image windows
for shape and position recognition.

成形わくの本体は二つの半円筒の鉄製容器であつて、側面に付属した数十本のボルトを締めることによって成形わくを円筒形にしたのちコンクリートを注入し、これが固化したあとボルトを緩めて成形された柱体を取出すという作業である。ボルトはその間隔が一定でなく、また横方向のフランジ部の奥まった位置にあり、成形わく回転用タイヤや補強用リブと隣接しているため、視覚装置などによる遠隔的検出が不可欠となる。

こうした理由により、上述の視覚装置を適用し、約 470 mm の位置から自立ボルトの支持部やリブを眺めるようにし、斜上方からこの部分に照明を当てた。第 11 図(a), (b)はこのようにして得たボルト、リブの映像である。タイヤは別の手段で簡単に検出できるので、視覚装置はこのボルトとリブを識別し位置を求める目的をもつ。

視覚装置の視野は水平方向に 120 mm、垂直方向に 90 mm と 1 ドラムの 120 mm が 320 組表におけるアーチ

なった。この値は油圧マニピュレータの先端にスプリングを介して取付けたインパクト・レンチのかみ合せ余裕内にあり、更にかみ合せ直前にスイッチを用いた触覚装置によりボルトの存在を再確認するとともに、より高い精度で位置の再測定を行なうので、自動ボルト締緩装置にとって十分なものとなっている。

形状認識の設定条件は、ボルトの支持部の横幅が 30 mm、リブの横幅が 9~12 mm であるから、例えばボルトに対して $h=32, m=40, n=32, \varepsilon_s=8$ 絵素とし、リブに対して $h=32, m=14, n=32, \varepsilon_s=12$ 絵素とすれば、27~33 mm の横幅をもつ物体がボルトとして、6~15 mm の横幅をもつ物体がリブとして認識されることになる。

なお、使用したビジョンの映像信号減衰率 λ は約 31 s^{-1} であるから、(3)式で定義した δ はボルトの場合 11.2、リブの場合 3.3~4.5 となる。従って、映像信号を 2 値化する条件として θ を 0.5 に選ぶとき、(6)式から導かれる物体像の横幅の減少分は、物体上で長さに換算してリブでも 0.10 mm 以下であり、ボルトの場合は更に小さいから全く無視できる。また(7)式から導かれる物体像中心の位置ずれは、同様に物体上で換算してボルトで 1.86 mm、リブで 1.82~1.83 mm だけ遅れる方向に生じることになる。これらの値自体は大きいが互いの差が僅少なため、前述の位置情報送出パルスの出力タイミングをずらすことにより十分な補正が可能である。

さて、撮像装置は対象物体に対し前進後退するようになっており、テーパを有する成形わくの通過につれて対象と撮像装置の距離が一定となるよう位置制御される。いまボルトが検出されると、その通過タイミング信号はデジタル遅延回路に入り、成形わくの移動方向に対し下流 1 m の所にあるマニピュレータを作動させ、ボルトに追従しながら触覚による確認動作とインパクト・レンチによる締緩作業を行なわせる。もしリブが検出されると、その通過タイミング信号は同じく遅延されたのち、マニピュレータに後退動作を行なわせ、インパクト・レンチがリブに衝突するのを避けるようにする。これにより平均的にボルト 1 本あたり 2.5 s の高速ボルト締緩作業が完全に無人作業として実現される。

置規模の儘小化を念頭に簡易かつ安定な視覚装置方式を開発した。特に積分形の演算を主体とした処理方式が映像雑音に対して極めて高い安定度を示すことから、4分割形の複数窓わくによる形状マッチング法と通過位置検出法を開発した。

この方式を電子回路で実現し、油圧マニピュレータと組合せて視覚と触覚を有する工業用知能ロボットとして構成し、視覚による外界対象の判断結果に応じて自動的に動作モードを変更させ、外界に適応した作業をさせるという試みを行ない好結果を得た。

この視覚装置における形状認識方式は複数個の窓わくの設定により、対象物体の代表的な部分を抽出するものであり、これらの数を増せば更に複雑な対象に対しても応用できるし、判定回路の複合化によって一層汎用性の高い工業用視覚装置の実現が期待できる。

また、ここでは位置が1次元内に規制され移動する対象を主体に考察したが、この移動方向と直角方向にずれる対象に対しても、ずれる方向に4分割した縦長の窓わくを発生し、かつこの窓わくの位置を浮動形にし、検出結果のフィードバックにより窓わく位置を制

御するよう構成すれば認識できるようになる。

このように、以上の視覚装置は単純な原理と構成によって、安定かつ確実な認識を行なえるという大きな特徴を持つものである。また、これまで研究室の域を出なかった知能ロボットから脱皮したハンド・アイ・システム実用機としての意義も大きいと思われる。

最後に、本研究開発の遂行にあたりご指導とご鞭撻を賜わった日立製作所中央研究所 渡辺所長、上妻副所長(当時)、沼倉部長(当時)、同社多賀工場 西宮工場長に、また認識アルゴリズムや装置開発にご協力いただいた同社中央研究所 池田貞弘、本池 順、同社多賀工場 磯尾 修、松岡慎二の諸氏に深甚の謝意を表わす。

(昭和50年8月5日受付)

文 献

- (1) 江尻・鳥野：電学誌 93, 986 (昭 48-11)
- (2) 他：同上 94, 771 (昭 49-9)
- (3) N. J. Nilsson: Proc. 1st IJCAI (1969) Washington D.C.
- (4) J. Feldman, et al.: Proc. 2nd IJCAI (1971) London
- (5) M. Ejiri, et al.: IEEE Trans. Computers C-21, 161(1972)
- (6) 鳥野、他：電子通信学会研究会資料 PRL 74-29 (昭 49-11)