

### 3次元画像入力装置とその利用について

梅木直子、森下明、沼崎俊一、土井美和子

(株)東芝 研究開発センター

我々は新しい画像入力デバイス **Motion Processor** を開発した。この **Motion Processor** はコンピュータへの簡単で、直感的な入力方法をユーザに提供するものである。**Motion Processor** は近赤外光を発光し、その反射光を撮像することによって、背景から切出された対象物の形状と動き、装置からの距離情報などをリアルタイムに取得することができる。これによってユーザは手振りや、身振りを入力することができる。我々は **Motion Processor** による直感操作体系の確立の一步として、また、従来のポインティングデバイスでは困難であった直感的な操作が可能となったことを実証するために、いくつかのデモアプリケーションを試作した。本文では **Motion Processor** の原理と応用アプリケーションについて述べる。

### The Application of a New 3D Image Input Device

Naoko Umeki, Akira Morishita, Shunichi Numazaki, Miwako Doi  
Toshiba Research and Development Center

We developed a new input device "**Motion Processor**". The **Motion Processor** presents an easy and intuitive input method to users. It throws infrared light to the object and receives the reflected light. So the **Motion Processor** can reject background and capture only the object's shape, motion and distance. Users can input the motion and the gesture to computers by using this device. We developed some applications for demonstration that this device is more useful for intuitive input. We will apply this device to the entertainment and edutainment applications. This paper describes the principle of the **Motion Processor** and some applications.

## 1. はじめに

現在パソコンは多くのオフィスに普及し、いよいよ家庭へと普及し始めている。しかし、まだまだその利用は難しく、買ってはみたものの、大して利用されていないのが現状である。その難しさの原因の一つが、入力手段が未だ限定されていることにある。現在の主流は、キーボード、マウスであるが、これらは修練を必要とし、初心者には敷居が高い。もっと日常生活でのコミュニケーション方法に近い、音声や身振り、手振りといったものを利用したいと以前から考えられてきた。音声入力最近、個人ユーザまでが利用できるような実用レベルのものが出来た。しかし身振り、手振りについては、未だ実用には至っていない。身振り、手振りの認識の研究は、様々になされてきている[1,2,4]。しかし、これらの研究では撮像した画像から、手などの対象物を背景から切出す事に処理時間が費やされている。この処理に多くの時間と計算量を費やしているため、リアルタイムに身振り、手振りを認識し、その結果を反映してアプリケーションを動作させるという目的を満たすには、問題が多かった。

そこで我々はそのような問題を回避すべく、従来の画像取得方法によらず、新しい方法で対象画像を取得するデバイス「**Motion Processor**」を開発した。

以下、**Motion Processor**の構成、**Motion Processor**で得られたデータを用いて認識処理を行うために開発した **Software Development Kit(SDK)**、さらに **Motion Processor**を用いたアプリケーションの例について述べていく。

## 2. **Motion Processor**の構成

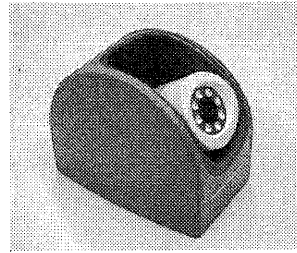


Fig.1 **Motion Processor** 試作機

**Motion Processor**は、光の反射の原理を用いて、背景画像を取り除くことで、対象物の画像のみを取り出す新規画像処理技術を採用した情報入力装置であり、対象物の形状や動き、装置からの距離情報などをリアルタイムに取得・認識することができる。

Fig.1は我々が試作した **Motion Processor**である。中央のレンズの周囲に、近赤外光を発するLEDを配している。またレンズの奥に受光部のカスタムセンサチップがある。

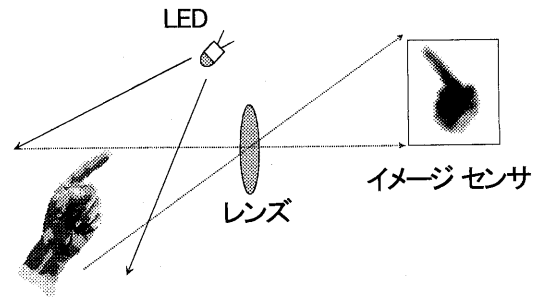


Fig.2 動作原理

Fig.2は **Motion Processor**の動作原理を示している。LEDから発光し、センサ部分が対象物からの反射光を受光する。

この様にして **Motion Processor**によって検出された手の画像の例をFig.3に示す。

図の様に、対象物である手の部分のみが検出されており、背景は検出されていないことが分かる。これは、背景からは僅かな光しか反射しないためである。この原理に

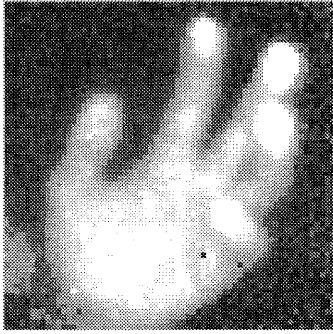


Fig.3 検出された手の画像

より、**Motion Processor** では、従来の画像処理の問題点であった、画像中からの対象物の切り出しを、低コストで容易に実現している。

$$I = \frac{k}{D^2} \quad (1)$$

I : 反射光の強度  
D : 物体との距離  
k : 定数

また、**Motion Processor** は、単に 2 次元画像だけでなく、3 次元情報を含んだ距離画像を取得することも可能である。反射光の強度は、式(1)のように、対象物までの距離の二乗に反比例して減衰する。

この性質を利用して、対象物からの反射光の強度を調べることで、対象物の距離画像が得られるという訳である。厳密に言えば、反射光の強度は、対象物体の表面の反射係数に依存するが、手のように全体を通して大体均一な物体であれば、反射光の強度をそのまま距離に対応させることも可能となる。Fig.4 に **Motion Processor** によって取得された手の距離画像の例を示す。

試作した **Motion Processor** の性能を Tab.1 に示す。コンピュータには、64x64 のデータアレイに 0 から 127 までの数値が格納されて送信される。

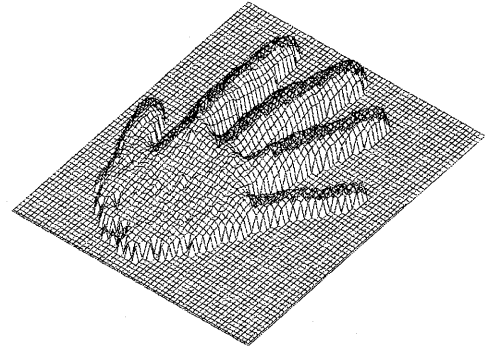


Fig.4 撮像された手の 3 次元表示

Tab.1 試作機の性能

分解能	64×64 pixels
距離深度	128 階調
撮像レンジ	30 ~ 90 cm
動作レート	30 ~ 50 frames / sec
本体サイズ	幅 90×高 120×奥行 140mm
対応 I/F	PCMCIA
対応 OS	Windows NT

### 3. 撮像画像による認識

我々は **Motion Processor** によって得られた撮像データから、対象物体の動きや形状を認識するためにライブラリを開発し、これを **Motion Processor SDK**(Software Development Kit)と呼んでいる。この中には、対象物体の重心位置や重心速度、面積、体積、FFT 演算、パターンマッチ、オプティカル・フローなどの演算を行う様々な API (Application Programming Interface)が含まれている。これらを用いることで、手の動きの追跡、形状の変化といったような様々な認識をリアルタイムに行うことが可能である。開発者は、これらの API を利用する事により、Visual Basic など簡単に **Motion Processor** を使ったアプリケーションを開発する事ができる。

次に、**Motion Processor** を利用したアプリケーションの例を示す。

#### 4. **Motion Processor** を利用したアプリケーション

**Motion Processor** により、従来のポインティングデバイスでは困難であった直感的な操作が可能となったことを実感できるように、また、直感操作体系の確立の一步として、いくつかのデモアプリケーションを試作した。

##### 4.1 バーチャルペット「いいこいいこ」

猫などのペットと遊ぶときには、頭をなげたりする。このような動作と同様に、画面上に表示されているキャラクタを、プレイヤーが手のジェスチャによってコントロールして遊ぶアプリケーションがバーチャルペット「いいこいいこ」である

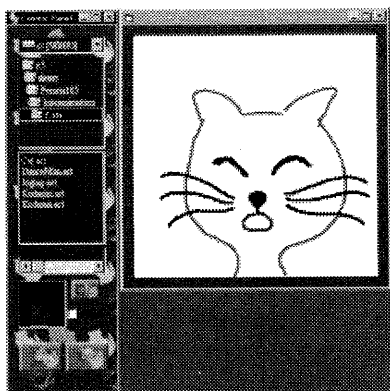


Fig.5 猫を撫でたところ

Fig.5 は、キャラクタの猫に対してプレイヤーが「撫ぜる」ように手を左右に振ると、猫が喜んでいる画面である。「撫ぜる」操作は、**Motion Processor SDK** の横方向の重心の位置や移動速度を演算する API により実現されている。

また、「叩く」ように手を前後に振ると、Fig.6 のように猫が怒るようになる。「叩

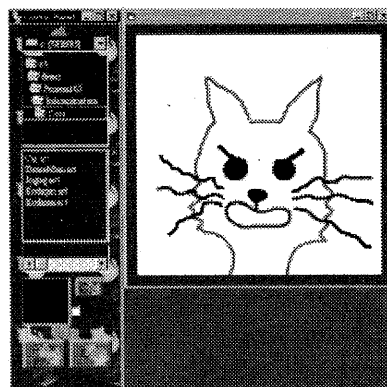


Fig.6 猫を叩いたところ

く」操作は、奥行き方向の重心移動速度により認識している。このように、複数の API の組み合わせによって、開発者が意図する動作の認識をさせ、その認識結果に対応した表示に変化させることができる。

従来のアプリケーションであれば、マウスによりメニューを選択して、デジタルペットを動かすという人工的な操作であった。これに対し、**Motion Processor** を使えば、プレイヤー自身の手の日常的な動作により、キャラクタとのより自然なコミュニケーションを楽しむことができる。

##### 4.2 バーチャル楽器「リズムセンサー」



Fig.7 バーチャル楽器を演奏している様子

打楽器など楽器演奏では、手が活躍する。同様に手の動きによって音を変化させて鳴らす「リズムセンサー」なるアプリケーション

ョンを試作した。「リズムセンサー」では、手の奥行き方向の移動速度や位置、移動方向などの演算結果を利用して、音の大きさ、種類、周波数などを変化させて鳴らす事ができる。このような動きにより音を扱う操作はさらに、子供向けに、体の動きに合わせて音が鳴ると同時に、絵が描けるといった、より創造的なアプリケーションに発展させることができる。

#### 4.3 フリースローゲーム

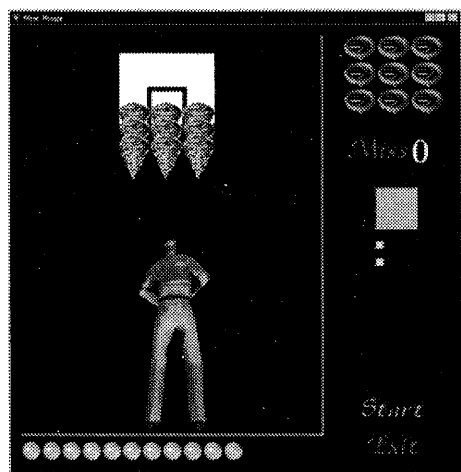


Fig.8 フリースローゲームの画面

これは某 TV 番組で有名な、バスケットフリースローゲーム「Nine Hoops」を、コンピュータ上で実現してみようと、試作したアプリケーションである。ボールを投げるとき、上と奥への強さ、方向といった複数のパラメータを、**Motion Processor** ならば一度に、しかも実世界と同様の動作で入力する事ができる。Fig.8 は、試作したフリースローゲームの画面である。

#### 4.4 じゃんけんゲーム

上述のアプリケーションでは手の動きを利用している。これに対し、手の動きと形状の認識の双方を行うことが必要なのが、じゃんけんゲームである。Fig.9 はじゃんけ

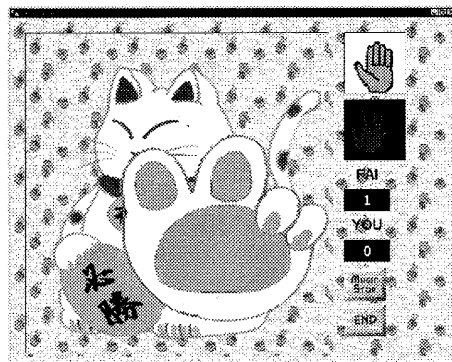


Fig.9 じゃんけんゲームの画面

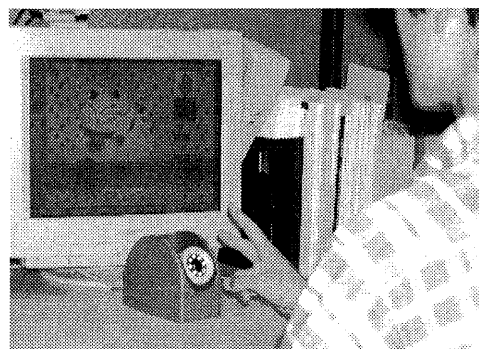


Fig.10 じゃんけんゲームをしている様子

んゲームの画面例で、中央の招き猫がじゃんけんの相手である。グー、チョキ、パーのいずれが出されたかの判定は、**Motion Processor SDK** におけるパターン認識の演算を用いている。また、プレイヤーの手の振りを認識してパソコンのキャラクタが「じゃん、けん、ぽん」と言うようになっている。従って、人間同士でじゃんけんをする場合と感覚的にほとんど差がないじゃんけんができる。

#### 4.5 両手を使って絵で遊ぶ「びよびよ」

実生活では、片手だけでなく両手を使った動作が多々ある。**Motion Processor** による操作でも両手を使うことが可能である。試作した両手操作デモ「びよびよ」では、両手を使ってビットマップイメージを自由にかつリアルタイムで変形して、遊ぶこと

ができる。1枚のビットマップイメージに対して、あらかじめ2つのコントロールポイントが設定されている。この点をそれぞれ両手で摘まんで引っ張るような感覚で操作すればよい。



Fig.11 びよびよの画面

Fig.11は、顔写真をひっぱってみたところの画面イメージである。効果音を組み合わせたりすると、デジタルカメラなどで撮像したイメージに対して、簡単に変形を加えて、遊ぶことができる。また、キャラクター全体が写ったイメージを用いれば、手で踊らしているような感覚になる。

このように、両手操作は、マウスやキーボードではできない多点制御であり、**Motion Processor**はこれに向いている。

## 5. おわりに

本論文では、我々が開発した新規3次元画像入力装置**Motion Processor**と、撮像画像データの演算処理を行う**Motion Processor SDK**と、その直感操作性を実証するために試作したデモアプリケーションについて紹介した。

**Motion Processor**は、物体の反射画像を撮像することで、形状や動きおよび物体までの距離情報を取得する新規入力装置である。これにより、従来の画像処理において

最大のネックであった物体の切り出しが、コンピュータに負荷をかける事なく、リアルタイムで行えるようになった。さらに、動作の認識を行うためのSDKを開発し、手の動作などの直感操作によるアプリケーションの開発環境を整えた。

今回紹介したアプリケーションは、エンターテインメントをターゲットにしたものばかりであるが、もちろん**Motion Processor**はこの分野だけでなく、手話認識、顔の認証、VR用の入力装置、監視装置、ロボットの眼、家電製品への組み込みなど、今後多方面に渡る応用が可能である。

今後は、この新規デバイスの性能向上、およびSDKの機能の充実を図っていく予定である。さらに**Motion Processor**による直感操作体系の確立をめざす。

## ● 参考文献

- [1] Cipolla, R., Okamoto, Y. and Kuno, Y. Qualitative visual interpretation of 3D hand gestures using motion parallax. *Proceedings of MVA '92*, pp.477-482, 1992.
- [2] 鹿毛裕史、三宅康也、久間和生、"人工網膜チップを用いたジェスチャ認識によるゲーム操作", 1997年電子情報通信学会総合大会、pp.364, 1997
- [3] Numazaki, S., Morishita, A., Umeki, N. et al. A Kinetic and 3D Image Input Device. *CHI 98 SUMMARY*, pp.237-238, 1998
- [4] 海老原一之、棚沢順、大谷淳、"赤外線カメラを用いた実時間人物全身像動き検出", 1996年電子情報通信学会基礎・境界ソサイエティ大会、pp.212, 1996