

PC上における3Dキャラクターの リアルタイムレンダリングでの再現性

弓田 貴彦, 片山 智博, 岩田 浩, 鷹取 利昌 (総合計画センター)

総合計画センター
〒650 神戸市中央区栄町通 6-1-21 神戸東邦生命ビル 5F

広く親しんでいただける手話教育システムの構築を目指して、試行錯誤を行っている。アニメーションでは、大量のデータを蓄積しておかねばならず、その作成労力も大きい。そこで、複数のキャラクターモデルと様々なモーションデータを作成しておき、その都度リアルタイムレンダリングを行うことを検討した。一般的な顧客各位の環境で使用できるよう、パーソナルコンピュータ上での構築を試行している。現時点では、なんとか手話は伝わる程度で、見映えは良くない。再現性の向上に努めている経過報告を行う。

Reappearancibility about Realtime Rendering of 3D Characters on PC

Takahiko Yumita, Tomohiro Katayama, Hiroshi Iwata, Toshimasa Takatori(SKC)

Sougou Keikaku Center
6-1-21, Sakaemachi-dori, chuo-ku, Kobe, Hyogo, 650 Japan

We are now constructing the Sign Language Education System which is friendly and generally. If the system uses CG animations, it is too many data that the general system is required. Then we thought that the system deals realtime rendering with several character models and many motion data. We can't use any WorkStaions and are now constructing on Personal Computers, because general customers can use PCs. Sign Language on the system can be recognized by audiences, but it has no good looking. We are now make efforts for better looking and better reappearance.

1 はじめに

我々は従来、土木用途のCGを業務の一つとしてきた。そのために用意してきたCG設備や得てきたノウハウを生かして、縁あって手話システム構築の研究が発生し、現在も進行中である。ただ、当初はCGキャラクター、手話ともにノウハウを持っておらず、それぞれで常識的なことから学ばねばならなかった。

本稿では、この手話システム構築上の経過や試行錯誤を紹介し、今後の展望について述べる。

2 コンセプトモデル

初期の段階では、まずコンセプトモデルとして、手話アニメーションを作成するところから取り組んだ。当初はキャラクターや手話に関するノウハウもなく、試行錯誤的要素が強かった。

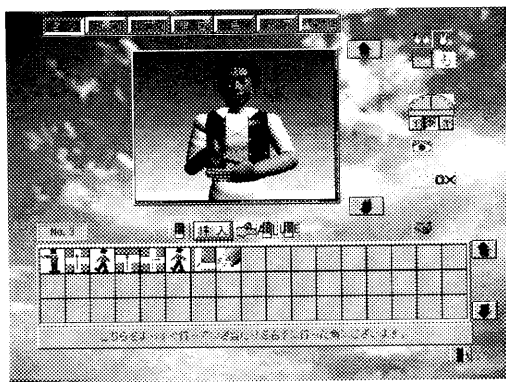


図1. コンセプトモデルでの手話アニメーション

2.1 コンセプトモデルの主な仕様

CGアニメーションによるシステムを構築した場合、自ずと限界があるのは当初より分かっていた。まずは顧客へ完成イメージを提供するため、コンセプトモデルとして制作した。

この版では、絵文字によるボタンを選ぶことに

よって文章を完成させ、CGアニメーションによって手話および指文字を表示するものであった。これは用意したアニメーションを、ボタン選択によって組み合わせていくものであり、予め用意できるアニメーションの量を考えれば、その限界は低いものであった。

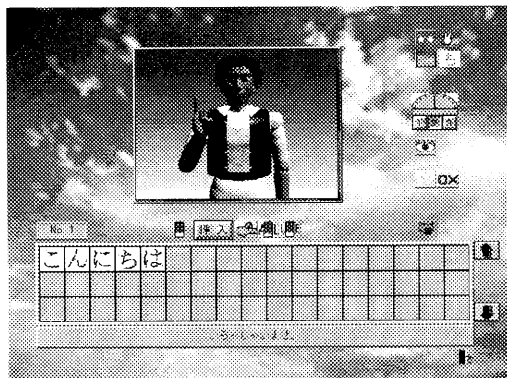


図2. コンセプトモデルでの指文字アニメーション

2.2 コンセプトモデルの評価

当初の予想通り、CGアニメーションによる手話表現は、すべてを賄うためには非常に多くのデータを蓄積する必要があった。それは、一般的な教育を行う上で、専用のコンピュータを導入しなければならないことを示し、これでは「広く親しまれるように」という主旨に沿わない。また、その広く一般に対応するデータ作成自体が容易でないとも言える。

逆に、この版では、最終目標となるシステムの質を示すことができた判断する。

3 テクニカルモデル

アニメーションやビデオによる、継ぎはぎの手話システムでは、自ずと限界が見えてくる。また、アニメーションではデータ容量が大きく、システム全体が大量の「辞書」的なものになってしまう。

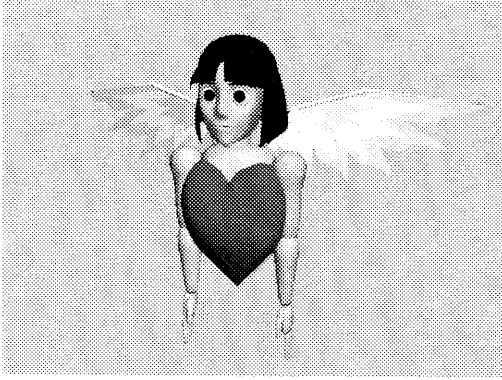


図3. 新たに作成したキャラクターイメージ

そこで、複数のキャラクターモデルとそのモーションデータを記録しておき、文章を構成する都度、手話モーションを構成、リアルタイムレンダリングを行うことを検討した。もちろん、ワークステーションレベルになってしまうと、広く親しんでいただくシステムとしては高価すぎる。あくまでも、既存のコンピュータを手話教育にも利用できるレベルを追求するため、パーソナルコンピュータレベルで稼動することを目標とした。

3.1 テクニカルモデルの主な仕様

コンセプトモデルの仕様を見直し、より扱いやすいものを目指した。

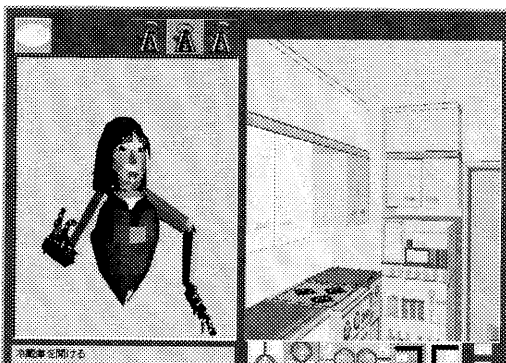


図4. テクニカルモデルでの手話レンダリング

文章構成部分をボタンではなく、ワークスルーを用いて、物や動作を選んでいく方式にした。これは、手話表現に必要なリアルタイムレンダリングを導入することに付随してできる機能を盛り込んだ、ということである。

文章を構成した後、3Dキャラクターが手話動作を行うのは、コンセプトモデルと同じであるが、これをリアルタイムレンダリング化した。この際、キャラクターは、コンセプトモデルの時よりも粗い表現となっている。

3.2 モーションキャプチャ

当然のことながら、3Dキャラクターモデルは、アニメーション作成時と同様にできる。テクニカルモデル作成時には、新たにデザインしたキャラクターを用いた。これは、コンセプトモデルよりも親しまれやすいキャラクターを、という配慮であった。

さて、モーションデータを採用する場合、現在ではいくつかの手法がある。我々が採用したのは、米 Motion Analysis 社の光学式モーションキャプチャシステムとサイバークローブであった。

手話を行う際の重要なモーションは手(掌, 腕)である。そのほか、手話では表現しきれない部分、伝わりにくい場合に有効活用されるのが、ある時は指文字であり、またある時は表情である。以上の範囲を網羅するためには上半身全体のモーションデータが必要となる。これら表情、上半身(腕を含む)、手(掌)は、それぞれ必要とするデータの密度が異なり、また、それぞれに最適なモーション採取方法も異なる。そこで、データをそれぞれ別個に採取し、連携を行うこととした。

表情、上半身(腕を含む)については、それぞれ光学式センサーを用いて、手(掌)については、サイバークローブを用いて、モーションデータを採取した。この際、米 Motion Analysis 社の御協力を得て、光学式センサー制御ソフトウェア "Eva" に、サイバークローブのデータ連携処理を行うよ

う、カスタマイズを行った。



図5. モーションキャプチャ機器

3.3 リアルタイムレンダリング

現在では、リアルタイムレンダリングを実現しているソフトウェアは、いくつか存在している。しかしながら、それらは様々な制約の下での実現である。また、本件のように3Dキャラクターを動作させることについては、アニメーションでも難しいと言われる表情の問題も生じる。特にリアルタイムで動作させるためには、データ量の制約もあり、表情を表現するのは簡単ではない、と考える。

まずは次の版をと制作した。

3.4 テクニカルモデルの評価

組み立てた文章から一応の手話表示が可能となり、その画面表示から意を汲むことまではできる程度となった。しかし、その手話動作は非常にぎこちない。手話を間違いないく伝える上で重要な補助となる表情の変化は、モーションデータを探っているにも関わらず、その見映えは実際の人間の表情にはほど遠く、それどころか表情に変化があるのかすら、よく見ないと分からない状態である。

4 将来展望

ここまでで記したように、PC上でのリアルタイムレンダリングは、まだまだ満足のいくレベルには完成していない。これから目指す方向性とは、どのようなものだろうか。

4.1 モーションデータの精度

テクニカルモデルにおける画面表示では、キャラクターの動作がどうもぎこちない。これには、いくつかの原因が考えられる。

- a) モーションデータの採取精度が低い
- b) モーションデータの修正が不十分
- c) リアルタイムレンダリングで、

いわゆる駒落ちを生じる

- d) Walk Throughを同時に動作可能にしている

これらのうち最も効いているのは、採取時もしくは修正後のモーションデータの精度で、手の震えなどの修正精度がさらに向上すれば、完成品の見映えが良くなるものと考えられる。現時点では、モーションキャプチャ機器の導入早々のデータを用いているため、オペレーターの技術的な慣れによっても、修正精度の向上を期待している。

4.2 モーションデータの修正作業時間

当初の見込みでは、アニメーションにするよりもモーションデータで蓄積した方が、辞書的部分の作成労力がはるかに小さいだろう、ということだった。しかし、モーションデータは採取よりも修正に長時間を要することが、経験的に分かった。オペレーターの慣れの問題はあろうが、現在のところ、5~10 words/dayといったところである。また、このオペレーターはワークステーションを操作できる技術が必要である。商業ベースで考えると、この高いコストの低減も一つの課題である。

4.3 表情の豊かさ

表情の表現については、考えられる課題を挙げておく。

まずは、リアルタイムレンダリングのためにデータを軽くしていることである。モデルデータを軽くすることによって、確かにリアルタイムへの寄与は大きいですが、逆に変化点を少なくしていることから、表情の再現が難しくなっている。

今後の課題として、顔の制御点を増やすことが一点である。

二点めとしては、モーションデータを忠実に反映できた場合でも、おそらく表情は豊かとは言えないであろう、ということである。人間の表情をフォトリアリスティックに再現すれば、表情豊かにできるであろうが、それは現在リアルタイムには不可能なので、どうしても漫画のアニメーションレベルのものにならざるを得ない。一般に漫画

のアニメーションでは、実際の人間の表情および顔の部分を、かなり誇張してストーリーを視覚的に理解させている。ところが、今回モデルに載せたモーションデータは、実際の人間のものであった。

リアルタイムレンダリングを行う上で、モデルデータを軽くすることは必須条件ゆえ、見映えを良くするためには、モーションデータの誇張を行わねばならないだろうと考えている。

5 結語

現在、我々の目指す手話教育システムは、まだまだ開発途上である。コンピュータの処理能力の向上は目覚しいが、それ以外の必要なノウハウを体得し、エンドユーザーのための手話教育システム構築に努めていく所存である。