

## 4. 仮想環境を実現するコンピュータプラットフォーム

A Computer Platform for Virtual Environment by Fumio KISHINO (Graduate School of Engineering, Osaka University).

岸野文郎<sup>1</sup>

<sup>1</sup> 大阪大学大学院工学研究科

### 1. はじめに

仮想環境とは、コンピュータによって合成された、現実と同様の機能を有する環境のことであり、合成品質はコンピュータの性能に大きく左右される。本章では、コンピュータプラットフォームに関し、仮想環境を実現する立場から解説を試みたい。したがって、コンピュータそのものについてはほかにも解説があるので、CPU、メモリなどの動向は概説に留め、仮想環境構築、とくに画像情報に関連した部分の技術を中心に解説し、さらにこれらの技術を適用した場合に実現される夢の部分も含めた将来展望について述べる。

### 2. コンピュータ技術の動向

#### 2.1 コンピュータの進展

デジタルコンピュータは1945年に誕生し、当初は弾道計算などの軍事目的として使われた。その後も大型コンピュータにより、主として科学技術計算として大学、企業の研究室で使われていた。その後、60年代から70年代にかけて、事務処理分野にも導入が図られ、大量の事務処理、データ処理を実施するため汎用大型コンピュータが使われていた。80年代に入り、文書処理にも使われ始めるとともに、ワークステーションが登場し、ダウンサイジング化が進展し、その流れは現在も続いているといえよう。90年代からは映像をも扱えるコンピュータも登場し、マルチメディア時代へと突入しつつある。

この流れは日本のコンピュータ市場の売り上げ比率をみるとよく分かり、91年には辛うじて50%を確保していたメインフレームが92年には50%を割り96年には30%程度になろうとしている。逆に91年には20%程度であったPCが

年々驚異的に増加し、96年には50%に達しようとしている。

技術の進展がこの流れを支えており、とくにCPUの処理速度の増大と、メモリの大容量化・低価格化に負うところが大きい。PCに使用されるCPU性能の傾向は後節で示すが、2.2年に2倍の割で増大している。これらは集積回路作成技術の進歩に依存しており、現在は0.8 $\mu$ mルールから、0.5 $\mu$ m、0.35 $\mu$ m、0.25 $\mu$ mへと進展しており、0.8 $\mu$ mから0.25 $\mu$ mになることにより性能向上比は約30倍になるといわれている。CPU処理速度の向上を目指した挑戦はスーパーコンピュータの分野でも行われており、70年代から5年で10倍の性能向上を達成しており、90年代になると多数のCPUを使う並列処理を導入することにより、5年で1000倍に性能が向上している<sup>1)</sup>。しかし仮想環境の構築にあたっては、パーソナルユースが基本と考えられ、PCの進展により注目すべきであろう。

半導体メモリを含む蓄積機能に関しても高機能化・低廉化が進展している。蓄積装置は記憶容量と転送速度で分類することができ、分類例を図-1に示す。半導体メモリの大容量化、低価格化はCPUと同様に集積回路技術の進展に大きく依存している。また仮想環境構築にあたっては、膨大な画像情報を蓄積する必要があるため、磁気ディスク、光ディスクの大容量化、高速化、低価格化が従来に増して重要となる。

#### 2.2 今後のコンピュータの展開

今後、コンピュータはどこまで進歩していくのであろうか。従来はハードは同じとして、プログラムによって異なる機能やサービスを実現していた。しかし集積回路の微細化技術は極限に近づきつつあるといわれながらも、 $10^{10} \sim 10^{11}$ の素子

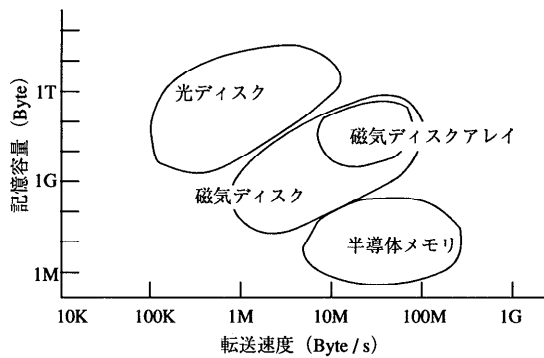


図-1 各種蓄積装置の位置づけ

を1チップに集積することが可能になりつつある。このような膨大な量の素子を検査して配線するのは不可能であり、自ら正常性をチェックし、結線も学習により自動的に進んでいくことを目指した研究も進められている。ハード的に人工脳を作ろうという挑戦的な試みである<sup>2)</sup>。

脳は学習によって処理方法を獲得する情報処理装置と捉えることができる。さらにその組合せで未経験な情報操作すら実現できるようになる。脳の学習には「情動」が深くかかわるとされており、細かな解析的な情報処理をとまわずに外的な変化に機敏に応答することが可能となる<sup>3)</sup>。このような人間らしい柔軟なコンピュータが実現されれば、仮想環境構築への適用も大いに考えられよう。

### 3. 仮想環境生成の動向

仮想環境を構築する場合に処理量が最も多いのは画像情報の構築である。画像情報を扱う技術としては、まず、入力画像から明瞭な画像を出力したり、画像を入力し、処理を施し、再び画像として出力する画像処理技術がある。次に、入力された画像に対してその画像が対象としているものの幾何学的、物理的属性あるいは物体の構成、さらにはその意味などのシーン・世界の記述を得る画像認識技術がある。さらに、シーンの記述や幾何学的あるいは物理的属性など、対象世界の記述が与えられたときそれに対応する画像を生成する画像生成技術がある。本章では、この中で仮想環境生成の要素技術として画像生成技術であるコンピュータグラフィクスと、画像符号化を中心とした画像処理技術の動向を述べ、コンピュータプラットフォームとしてどのようなものが適用可能かを解説する。画像認識技術については具体例の章で触

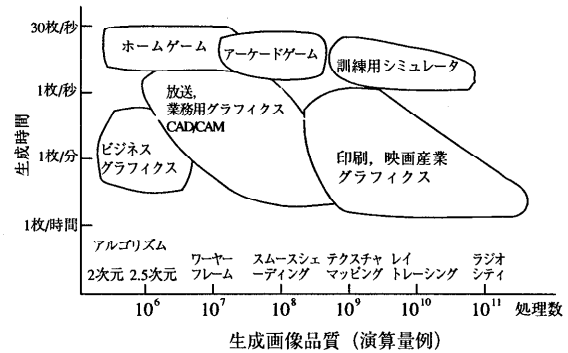


図-2 画像生成技術の評価

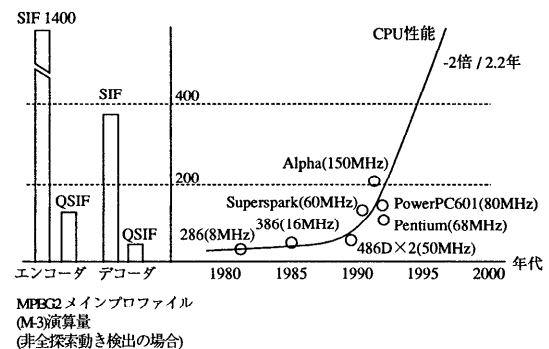


図-3 RISC性能とMPEG2エンコーダ処理量

れることにする。

#### 3.1 コンピュータグラフィクス(CG)

CG技術の初期時代に特筆すべきものとして、1963年にSutherland(MIT)がCRT上に描画する“SKETCHPAD”を発表している。これは稜線のみからなる簡単な図形であるが、ユーザの視点の位置にしたがって見え方が変化し、インタラクティブCGのはじめであると同時に、バーチャルリアリティの先駆けとしても知られており、CGの最初からこのような考えが芽生えていたのは興味深い事実である。その後、コンピュータの演算能力の進歩とともに画像生成技術も発展してきた。60年代は“SKETCHPAD”もそうであったように対象物を多面体で近似し、その稜線を表示する針金細工的なワイヤフレームモデルが主流であった。70年代に入ると物体の表面を表示し、さらに光を考慮した陰影画像を生成するサーフェースモデルの時代となった。80年代には物体の容積などを考慮したソリッドモデルの時代となり、光に関する計算アルゴリズムもレイトレーシング、ラジオシティなど種々考案され非常に現実感のある画像が生成できるようになってきた。90

年代になると、これらの現実感のある画像をリアルタイムで生成する時代になってきたといえよう<sup>4)</sup>。

CGで話題になったのは、映画のターミネータ2、ジュラシックパークでロボットから人間に変身するシーン、恐竜が疾走するシーン、などであり、これらはすべてCG技術を用いて作成されている。CG技術はモデリングとレンダリングに大別され、前者は仮想物体を多数の多角形(ポリゴンという)や曲面で形状を記述するとともに、表面の色や材質を定義するもので、後者はある視点

から眺めた画像として描いて表示するものである。映画の場合では、模型を元に詳細なモデルを作り、実際の人物、動物の動きを参考にして、1枚1枚手作業で作成しており、数100台の最新鋭のグラフィックスワークステーションを使用し、人手と時間をかけている。

画像生成時間と生成された画像の品質との関係を簡略化して概念的に図示したものを図-2に示す<sup>5)</sup>。映画の世界で実画像と見誤るぐらいリアルな画像を生成するには現状では実時間は不可能である。仮想環境生成、ヒューマンインタフェース



図-4 バーチャル歌舞伎システム

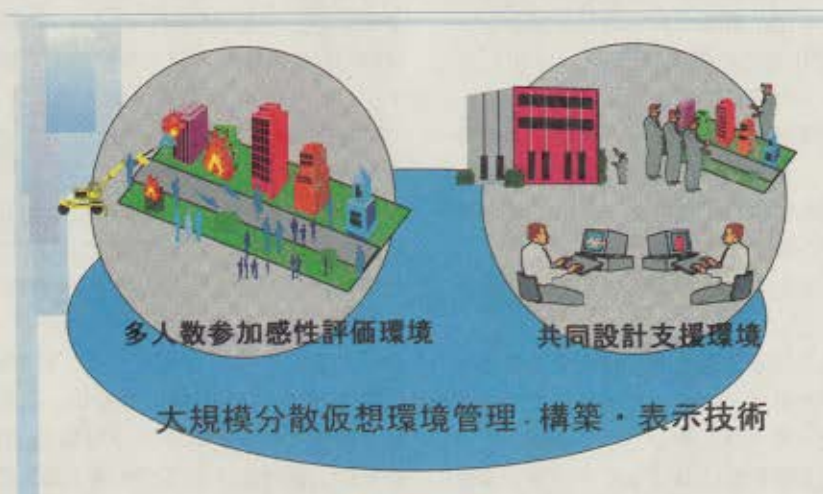


図-5 都市環境ヒューマンメディアのイメージ

などを考慮するとリアルタイム生成が必須となる。画質を犠牲にしてリアルタイム性を追及したのが、ゲーム機器に使用されるCGであり、最近では画質の向上もはかられている。1～2万円で購入できるのは驚異的であるといえる。訓練用シミュレータ、前述の映画用グラフィクスは画質を追及しているが非常に高価で億単位の価格となる。業務用のグラフィクスは100万から数千万円程度であろう。

高画質のCGをリアルタイムで生成するにはどの程度のCPU処理量が必要かを紹介する。ポリゴンによる画像生成では、数100MOPS～数10GOPSが必要となる。ここで、MOPSはMega Operation Per Secondであり、1MIPSは2～3MOPSとして換算する場合が多い。この場合、表示画素数は720×480を想定すると総ポリゴン面積を画面サイズと仮定しても数100MOPS必要となり、ポリゴン面積が増大するとさらに大きな処理量が必要となる。さらにレイトレーシングでは膨大な処理量となり、1秒間に30枚のリアルタイム画像生成を実現するためには約1TOPSの処理量が必要であり、テクスチャマッピング処理を加えるとさらに1桁必要となる。スーパーコンピュータの処理量は最近では5GFLOPS以上といわれており、超並列型では数100GFLOPSのものもある。レイトレーシングで生成しようとする超並列のスーパーコンピュータで辛うじてリアルタイム生成が可能となり、いかに現実的でないかご理解いただければ。なおマイクロプロセッサは最高速のものでも数100MIPS程度であり、現状ではリアルタイム化が困難であることが分かる。

### 3.2 画像符号化などの画像処理

インターネットの爆発的増大により、インターネット上で仮想環境を構築しようとする試みが多くなされている。実写画像を使用する例が多く、情報量の多い動画画像信号を圧縮伝送するための画像符号化技術が今後ますます重要となろう。ここでは画像符号化を中心とする画像処理について述べる。

従来専用装置で行ってきたテレビ画像品質並みの動画画像処理が汎用アーキテクチャや標準のインタフェースを使ってソフトウェアでできるようになってきた。画像符号化におけるリアルタイム処理について考察する。対象とする映像信号はスタ

ジオ規格である4:2:2デジタルコンポーネント信号である。ここで4は13.5MHzのサンプリング周波数を意味し、輝度信号はこの周波数でサンプリングし、2つの色信号はこの1/2の周波数でサンプリングすることを意味する。画像符号化の歴史は古く、画像処理であるフィルタによる雑音除去から始まり、隣接画素間の差分の符号化、フレーム間差分の符号化へと進展し、最近ではMPEG標準化で大きく花が開いた周波数領域変換の一種である離散コサイン変換(DCT)が主流となっている。DCTは74年に提案されていたが、当初処理量が大きく、リアルタイムでは動作できなかった。エレクトロニクス技術の進展に支えられ、最初専用装置で装置化され、最近では汎用装置とソフトウェアでリアルタイムで動作できるようになってきた。

デコーダLSIの製品化は第3フェーズといわれている。第1フェーズは開発が主体であり、DSPタイプが主流であった。0.5μmルールを用いて量産を指向した第2フェーズを経て、第3フェーズ(現在)は装置に完全に合わせ込んだチップセット開発の時代となっている<sup>8)</sup>。

エンコーダは後述するように処理量が多く、経済的な実現が困難であったが、MPEG2符号化の全機能を1枚のPCボードで実現したワンボード・エンコーダが商品化されるようになった<sup>6)</sup>。処理量の大半を占めるのが動きベクトル検出であり、基本的な全探索手法を使うと10TOPSの処理量となり、0.5～0.35μm技術ではLSI化は不可能となる。階層的探索法を採用することによりLSI化を実現している。汎用のCPUでMPEG2の処理が可能かを検討した例を図-3に示す。QSIF(通常のSIFの画素数の1/4)に対しては対応可能であるが、SIFに対してはデコーダでも対応不可能なことが分かる。

しかしCPU性能は2.2年で2倍になっているとのことであり、近い将来にはPCでもMPEG2の符号化が実現されることが期待される。

前節でCG生成について述べたが、画像符号化の処理量についても考察する。MPEG2の一般的な解像度は720×480であり、通常のテレビ信号と同様にフレームレートは毎秒30フレームを符号化、復号化する場合を考える。符号化の処理においては、時間方向の処理である動きベクトル

検出に多くの処理演算が必要であり、全体の約9割が必要となる。空間的な処理であるDCTには1割弱の処理量を占めるのみである。全体の処理量は約35GOPSである。一方、復号化では空間的な処理であるDCT処理に多く使われ、全体で1.8GOPSのうち約7割強が必要となる。なお、MPEG1では符号化に数GOPSが必要となり、復号化では数100MOPSが必要となる<sup>7)</sup>。

画像符号化は単に圧縮するだけでなく画像処理と組み合わせて新しい展開が期待される。圧縮データと画像処理を組み合わせることにより、たとえば映像のカット点検出にMPEG2圧縮データ(動き予測誤差など)を使うことにより実現できる。フレームやフィールドの相関で技術の粋を尽くしたアルゴリズムであり、圧縮データは情報の宝庫といえることができる。ほかの例を次節でも紹介する。

前節で触れた集積回路技術の進展は画像処理技術にも多大の影響を与えており、従来のLSI化技術ではテレビ放送品質しか扱えないが、0.25  $\mu\text{m}$  ルールが実現できればHDTVを、0.15  $\mu\text{m}$  ルールになれば2048  $\times$  2048画素のSHDを扱うことが可能となり、医療診断や電子博物館などへの適用が期待される。

現在は多くのタスクの1つとして、マイクロプロセッサがソフトウェアデコードで画像のフルモーション再生機能を有することが当面の目標となっている。また、現在精力的に推進されているのはメディアプロセッサ(MP)とマルチメディア命令を搭載したハイエンドなマイクロプロセッサであるスーパープロセッサ(SP)であり、fax/modem, MPEG2 decode, H.261 ビデオ会議システムなどを搭載したマルチメディア対応のDSPであるMPは2~10GOPSの性能を有する<sup>8)</sup>。

このようにマイクロプロセッサの能力が向上すれば、ウェーブレット変換、ベクトル量子化、フラクタルなど非標準の符号化方式であってもデコードソフトをインターネットで配信するかコンテンツの頭に付与すればよく、新たな展開が期待されることとなる。

#### 4. 仮想環境の適用例

以上、画像処理、CG生成の処理量について述べ、コンピュータ処理能力の増大によりこれらの

処理がリアルタイムで可能になりつつあることを紹介した。これらを使ってどのような応用へ展開できるかを、筆者の専門分野であるバーチャルリアリティ関連を中心に紹介する。

##### 4.1 人物像の実時間認識・生成

人物像を対象に最新の研究成果を紹介する。人物像をCG表示するためには、あらかじめ人物の形状をポリゴンデータとして、また皮膚の色、着物の柄など(テクスチャーという)のデータを蓄積しておく。顔の表情の変化、手足などの人体の動きをリアルタイムで検出し、このデータを基にあらかじめ蓄積されている人物モデルを動かして表示すればリアルタイムで人物像をCG生成することができる。したがって画像処理・認識技術と画像生成技術を統合したシステムととらえることもできる。人物モデルとして実際に遠隔地にいる人のモデルを用いれば、仮想空間内を介して臨場感通信会議ができる<sup>9)</sup>。モデルとして歌舞伎役者を用いるものが図-4に紹介する「バーチャル歌舞伎」である<sup>10)</sup>。全身を赤外線カメラで撮り、顔、手足の部分をリアルタイムで認識し、それに合わせて人物モデルを動かす。顔の表情については顔の前に取り付けられたカメラから得られる画像から、目、口などの変化が激しい領域を選び、この部分の周波数領域上の特徴として顔の表情変化を検出する<sup>11)</sup>。従来は、リアルタイムで検出するためには、光ケーブルが縫込まれたボディースーツを着て、顔に動きを検出するマークを貼り付ける必要があり、ヒューマンインタフェース上好ましくなかった。周波数領域出の特徴はMPEGで使われるDCT(離散コサイン変換)を適用することができ、これらは前章で述べたようにリアルタイム化が可能となり、写真に示すことが実現できるようになった。この方法を用いれば、カメラを取り付けるヘルメットを被る必要はあるが、マーク、特殊スーツは不要で、人に優しいシステムにすることができる。

この技術を用いると人物のモデルパラメータのみを送信すればよく、インターネットでリアリティの高い仮想環境を構築できることが可能となる。

##### 4.2 自然環境を対象とした仮想環境構築

上述の人物像を対象とすると、3次元モデル化は比較的容易であるが、都市全体、あるいは野山

などの自然環境を対象として仮想環境を構築しようとするコンピュータ内でモデル化するのは存在する物体の数が多く、また形状が複雑であるため、多大の労力を必要とする。したがって対象とする実世界の環境を計算機に入力することが不可欠となる。3次元モデル化が困難であるため、3次元形状を復元せず、複数の既存画像を用いて仮想環境を構築することも提案されている<sup>12)</sup>。視点位置ごとに用意した画像を切り替えて表示したり、視線方向のみの変化は視点周囲の映像を記録しておき、切り出しにより対処することによりインタラクティブ性を実現している。また連続した視点の移動に対しては、モーフィングにより補間画像を生成することにより実空間と同一の機能を実現しようとしている。

3次元構造情報を獲得するため、複数のカメラ画像を用いる手法が検討されている。CMUのVirtualized Realityのプロジェクトであり、マルチベースラインステレオの手法を用いて動画像からの3次元構造情報の復元に関する研究を行っている<sup>13)</sup>。

実環境にコンピュータ内で作成した仮想環境を重畳表示しようとする研究も盛んである。オーグメンテッドリアリティと呼ばれており、この場合、実物の上にCG像を重畳表示すると、本来は物体表面の奥にあるものが、表面上に表示されることになり実際の感覚と異なることになる。これを避けるためには、実物体もコンピュータ内に取り込んで奥行き関係を計算し、CG像が実物体に隠される場合はその部分は表示しないようにする必要がある。

このような技術が確立されれば、都市全体をコンピュータ内に取り込み、都市設計においてその中の一部のビルを立て直したとき全体の景観はどのようなになるかのシミュレーションが可能となる。さらには、大震災が発生したときの想定被害や、避難シミュレーションへの応用も期待される。

具体的な検討例としてヒューマンメディアプロジェクトで進められている都市環境ヒューマンメディアを紹介する<sup>14)</sup>。本プロジェクトでは都市環境および都市を構成する工業製品などの設計および評価を、複数の異なる分野の専門家が協調できるプロトタイプ(共同設計支援環境)、およびこのプロトタイプにより構築される仮想都市環境を

多数の一般市民が同時に、相互にコミュニケーションを行いながら体験可能なシステムのプロトタイプの構築、ならびにこれらのシステムの構築に必要な基盤技術である大規模分散仮想環境の管理・構築・表示技術の確立を目指している。図-5に都市環境ヒューマンメディアのイメージを示す。

## 5. む す び

コンピュータプラットホームに関し、画像情報を中心とした仮想環境を実現する立場から解説を試みた。仮想環境を構築するためには音情報などほかのメディアも重要な位置を占めるが、処理量は画像情報が最も多いためである。

21世紀は高度情報社会が到来するといわれており、インターネットの爆発的普及とコンピュータの飛躍的進歩に支えられ、社会生活がネットワークを介したデジタル情報を中心に営まれることが想定される。このような社会において、すべての利用者にとって優しく、使いやすい仮想環境の構築が待たれる。

今回の解説につきましては、各種の講演、ディスカッションで得られた情報を参考にさせていただきました。日頃、ご指導、ご議論いただく皆様に深謝いたします。

## 参 考 文 献

- 1) 島崎：スーパーコンピューティング応用の現状と将来, 情報処理, Vol.36, No.2, p.125(1995).
- 2) 下原：人工生命(II)-情報処理への応用と展望-, テレビジョン学会誌, Vol.50, No.11, pp.1746-1754(1996).
- 3) 松本, 市川：脳におけるデータ処理とエレクトロニクス, 電子情報通信学会誌, Vol.79, No.11, pp.1060-1065(1996).
- 4) 中嶋：CGの基礎論, テレビジョン学会誌, Vol.46, No.1, pp.49-57(1992).
- 5) 鷺島, 西澤, 浅原：並列図形処理, コロナ社.
- 6) 山内, 泉岡, 谷中, 田代, 伊藤, 小寺：MPEG2パソコンボード, 情報処理学会研究会資料, AVS-8-7(1995).
- 7) 藤原, 大久保：インターネット時代の画像圧縮技術, アスキー, p.44(1996).
- 8) 岸本, 山内：動画像処理のためのエレクトロニクス, 電子情報通信学会誌, Vol.79, No.11, pp.1066-1074(1996).
- 9) 岸野, 宮里：人間主体の知的通信システムを目指して—臨場感通信会議—, 信学技報, MVE95-44, pp.1-8(1995).

- 10) Ohya, J. et al : Virtual Kabuki Theater :Towards the Realization of Human Metamorphosis Systems, RO-MAN '96, pp. 416-421 (1996).
- 11) 海老原, 大谷, 岸野 : 臨場感通信会議のための実時間表情検出, テレビジョン学会誌, Vol.50, No.10, pp1497-1506 (1996).
- 12) Eric Chen, S. : Quick Time VR -An Image-Based Approach to Virtual Environment Navigation, Proc. SIGGRAPH 95 (1995).
- 13) Virtualized Reality Home Page: <http://www.cs.cmu.edu/afs/cs/project/VirtualizedR/www/>
- 14) ヒューマンメディア先導研究 : <http://www.ctl.go.jp/ctl/taiwa/HumanMedia/>  
(平成9年3月6日受付)



岸野 文郎

昭和44年名古屋工業大学電子工学科卒業。昭和46年同大学院修士課程修了。同年日本電信電話公社(現NTT)電気通信研究所入所。同ヒューマンインタフェース研究所において、高速・広帯域通信システムの研究・実用化、画像処理の研究に従事。平成元年国際電気通信基礎技術研究所(ATR)に出向。ATR通信システム研究所知能処理研究室室長。臨場感通信会議、マルチモーダルインタフェースなどの研究に従事。ATR知能映像通信研究所を経て、平成8年7月より大阪大学大学院工学研究科教授。電子情報通信学会、映像情報メディア学会、人工知能学会、日本VR学会各会員。 e-mail: kishino@ise.eng.osaka-u.ac.jp