

2

Conceitos Básicos

Este capítulo apresenta em linhas gerais os problemas visuais mais comuns e, de modo mais específico, discute a atuação dos mecanismos de percepção humana especialmente na ausência da visão. Esta discussão está baseada em trabalhos científicos da área da psicologia de percepção e será de grande relevância para a especificação de técnicas não-visuais de interação voltadas especificamente para os deficientes visuais.

2.1 Deficiências Visuais

Embora este assunto esteja fora do escopo da tese, é importante que seja entendido o que é a deficiência visual, pois esse termo será utilizado com frequência neste trabalho. Um indivíduo é considerado deficiente visual quando, mesmo com correções óticas (óculos e lentes de contato) ou cirúrgicas, há um severo comprometimento de sua acuidade visual, isto é, capacidade de perceber detalhes de uma imagem, ou comprometimento de seu campo visual, isto é, capacidade de enxergar em várias direções (Lázaro, 2002). Outras formas de deficiência visual, tais como, dificuldades de percepção de cores, sensibilidade a contrastes e problemas de coordenação do movimento dos olhos também podem comprometer a capacidade visual (AOA, 2002a). Entretanto, esses aspectos não serão abordados neste trabalho.

O grau de deficiência visual pode variar desde cegueira total (ou cegueira funcional), onde o indivíduo não possui nenhuma sensibilidade luminosa, cegueira parcial (ou cegueira legal ou profissional) e visão subnormal, onde o indivíduo possui razoável resíduo visual (Congresso Brasileiro de Prevenção da Cegueira, 1980).

Pode-se citar ainda uma outra classificação do grau de deficiência visual, mais informal e bastante utilizada em práticas esportivas. Nesta classificação, os indivíduos podem ser B1 (totalmente ou praticamente cegos), B2 (pouco resíduo visual) ou B3 (bastante resíduo visual) (Conde, 2002). Por ser mais simples e de

uso geral, esta metodologia foi adotada para classificar os usuários nos testes realizados com o protótipo desenvolvido neste trabalho.

Segundo a Organização Mundial de Saúde, cerca de 1% da população mundial apresenta algum tipo de deficiência visual. Mais de 90% encontram-se nos países em desenvolvimento. Dois terços dos cegos encontram-se na Índia, China e em países africanos. Apenas um terço são homens. Nos países desenvolvidos, a população com deficiência visual é composta por cerca de 5% de crianças, enquanto os idosos são 75% desse contingente. (Entre Amigos, 2002; Folha de São Paulo, 2001). Dados estatísticos sobre a cegueira nos EEUU podem ser encontrados em (AFB, 2001). Logo, conclui-se que a cegueira é um problema que tipicamente atinge países sub-desenvolvidos ou em desenvolvimento, como é o caso do Brasil. Isto reforça ainda mais a relevância do tema.

As principais doenças e problemas relacionados com a visão são: catarata, glaucoma, retinose pigmentar, atrofia do nervo ótico e descolamento de retina (AOA, 2002b; Cerejo, 2002; RNIB, 2002a).

Recentemente tem-se notado uma tendência mundial em se desenvolver produtos que possam ser utilizados por uma parcela mais ampla possível da população. Esta preocupação é justificada considerando, por exemplo, o aumento da expectativa de vida na maioria dos países. Ao envelhecer, cresce sensivelmente a probabilidade de um indivíduo sofrer sérias restrições funcionais, tais como perda de visão, audição e capacidade cognitiva e motora (Vanderheinden, 2000).

Este tipo de enfoque deve, portanto, levar em consideração diversos tipos de restrições e atender a uma gama muito ampla de requisitos. Por este motivo, optou-se por focar esta pesquisa no estudo de interfaces para usuários cegos totais. Questões relacionadas com usuários de visão subnormal também não serão abordados aqui, pois tais indivíduos utilizam outros métodos, citados no capítulo anterior, que os permitem trabalhar com os atuais modelos de interface ou melhor aproveitar o resíduo visual que possuem.

Portanto, outras deficiências, além da cegueira, estão fora do escopo desta tese. Entretanto, existem vários trabalhos que tratam diretamente dessas questões (Gregor & Newell, 2000; Jaffe, 1994; Keates et al., 2000a; Keates et al., 2000b; Keates et al., 2000c; McLaughlin et al., 2001; Rowan et al., 2000; Smith et al., 1996; Trewin, 1996; Vanderheinden 2000).

2.2 Modelo Mental

Sabe-se que um indivíduo completamente cego, quando bem treinado, locomove-se facilmente sem precisar de auxílio em um ambiente conhecido (Smith, 2001). Embora não dispondo do canal visual, os deficientes visuais adquirem conhecimento do ambiente à sua volta utilizando outros mecanismos sensoriais. Neste caso, o ambiente é explorado principalmente através do tato, buscando propriedades que possam ser sentidas: forma, tamanho, textura, posição, disposição espacial etc. Outras propriedades, tais como cor, brilho e transparência, são ignoradas, pois dependem exclusivamente da visão. Essa exploração, ainda que bastante trabalhosa, ineficiente e demorada (quando comparada com a exploração visual), permite um deficiente visual ter boa compreensão do ambiente.

As observações acima levam psicólogos a discutirem a hipótese de indivíduos (videntes ou não) construírem um *modelo mental* do ambiente onde se encontram (Kurze, 1996). Os defensores dessa teoria acreditam que, baseados em suas experiências sensíveis, todos constroem mentalmente uma estrutura espacial de dados acoplada a um conjunto de possíveis ações sobre a mesma (ver Figura 2.1).

Caso uma determinada ação torne a estrutura inconsistente, ela é automaticamente reformulada até tornar-se consistente novamente. Por exemplo, se o deficiente visual espera encontrar uma parede em determinado local e, ao empurrá-la, sente algo se mover, o indivíduo assume que em tal local há uma porta ou janela. A seguir, atualiza seu modelo mental de forma a corrigir o erro na estrutura mental de dados.

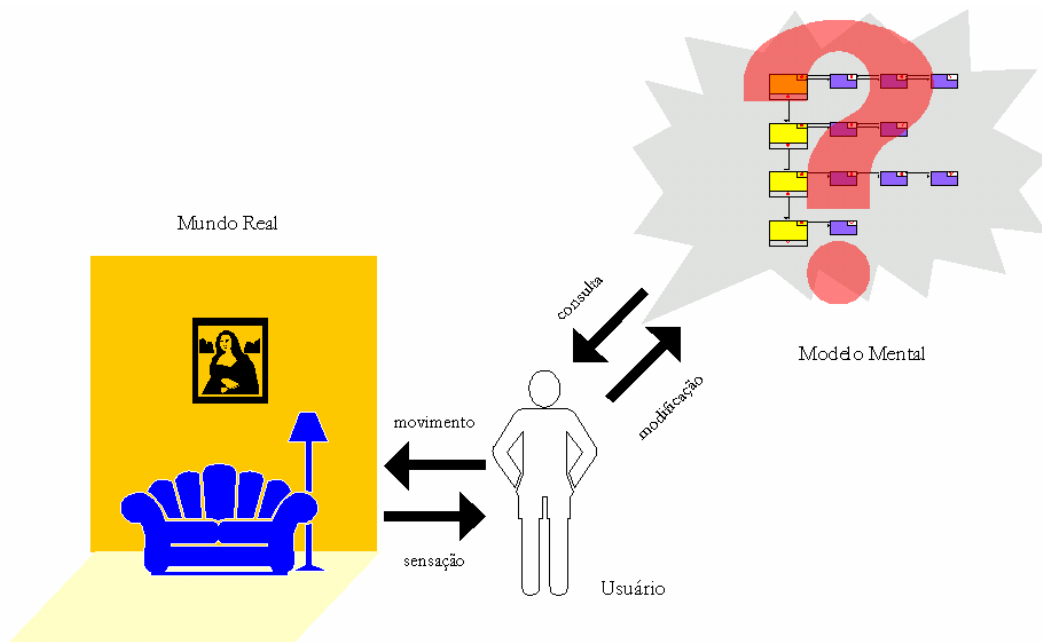


Figura 2.1: Modelo Mental

Preece et al. (1994) afirma que as pessoas constroem modelos mentais do mundo para dar suporte a suas ações executadas externamente. Confirmam também a idéia de que um modelo mental representa, entre outras informações, a disposição relativa de um conjunto de objetos no mundo. Portanto, está relacionado com o conceito de imagem, porém de forma funcionalmente distinta: uma imagem é a representação estática de uma cena, enquanto no modelo mental é dinâmica.

Pressupondo a existência de um modelo mental, Kurze (1996) define uma função de percepção e sua inversa, função de ação, conforme a Figura 2.2. Para um indivíduo vidente, a sensação é quase totalmente baseada na visão. Entretanto, para um cego ela é baseada principalmente no tato e na audição.

$$f_{\text{percepção}} : \text{Mundo Real} \xrightarrow{\text{sensação}} \text{Modelo Mental}$$

(a) percepção

$$f_{\text{ação}} : \text{Modelo Mental} \xrightarrow{\text{movimento}} \text{Mundo Real}$$

(b) ação

Figura 2.2: Funções de *percepção* e sua inversa *ação*

Imaginando que, em um ambiente familiar, o deficiente visual realiza praticamente as mesmas tarefas que um vidente, Kurze afirma que:

- O modelo mental de ambos deve ser completo, ou seja, contém todas as informações relevantes a respeito do ambiente;
- A função de ação ($f_{ação}$) para videntes e deficientes visuais é aproximadamente a mesma;

Neste caso, o autor conclui então que, guardada as devidas proporções, o modelo mental de ambos é aproximadamente o mesmo. A diferença está no método utilizado para construir o modelo: a função de percepção ($f_{percepção}$) dos videntes está baseada na sensação visual, enquanto a dos deficientes visuais baseia-se na sensação tátil e auditiva.

Pode-se ter uma indicação de que a conclusão acima é verdadeira perguntando a um deficiente visual como chegar a um determinado local. Certamente sua resposta será muito semelhante a resposta de um vidente, podendo conter até um nível maior de detalhe. Vale lembrar, entretanto, que este exemplo não comprova definitivamente a semelhança entre os modelos mentais. O uso da linguagem, por ser linear, não é a melhor mídia para demonstrar a hipótese. Nesse caso, seria mais adequada uma experiência não-linguística.

Não há dúvida que, entre deficientes visuais e videntes, há grandes diferenças no mecanismo de aquisição de conhecimento do mundo real. Entretanto, o mesmo questionamento poderia ser feito entre os deficientes visuais. Certamente, esse mecanismo também pode sofrer alterações. Uma das diferenças mais importantes pode estar relacionada com o fato do indivíduo ter nascido cego ou ter ficado cego. Embora não seja objetivo deste trabalho discutir com profundidade essa questão, ela pode ter implicações neste trabalho. Portanto, este fato será visto com mais cuidado ao ser analisado os testes com usuários. Isso justifica classificá-los segundo o tempo de cegueira.

2.3 Percepção Háptica

O termo “háptico” está diretamente associado ao sentido do tato. No ser humano, este sentido possui dois componentes independentes: cutâneo e cinético.

O componente cutâneo está ligado aos sensores localizados na superfície da pele. São responsáveis por sensações tais como pressão, temperatura, vibração e dor. O componente cinético está ligado aos sensores localizados nos músculos, tendões e juntas. São responsáveis por sensações tais como movimento e força (Oakley et al., 2000). Ao tocar e mover um objeto qualquer, o indivíduo terá diversas sensações que serão convertidas pelo cérebro em informação cutânea e cinética (Klatzky & Lederman, 2000). Isto permitirá ao indivíduo inferir propriedades deste objeto, tais como textura (de acordo com a análise da informação cutânea) e características gerais, tais como peso e forma (de acordo com a análise da informação cinética).

Uma prova irrefutável da capacidade de processamento temporal e espacial do tato humano é existência de um mecanismo de comunicação para surdo-cegos, conhecido por Tadoma (Reed, 1996). Contanto exclusivamente com o sentido do tato, um indivíduo surdo-cego é capaz de entender o que uma outra pessoa está falando colocando suas mãos sobre a face do interlocutor, percebendo assim vibrações e movimentos articulatórios dos lábios e maxilares.

Klatzky et al. (1993b) realizaram uma experiência simples que consistia em mostrar para uma pessoa vidente um par de objetos e, a seguir, perguntá-la qual era o maior, o mais pesado etc. Para as propriedades geométricas, tais como forma, área e volume, as pessoas tocavam nos objetos apenas para trazê-los mais próximos de si. Entretanto, para as propriedades materiais, tais como textura, rigidez e peso etc, o tato era efetivamente utilizado em detrimento da visão. Este resultado demonstra a natural tendência do ser humano utilizar a visão para perceber propriedades geométricas e o tato para propriedades materiais. Portanto, em um ambiente virtual, para que um deficiente visual possa construir um modelo mental adequado de objetos, torna-se particularmente relevante mecanismos que possibilitem a tradução de propriedades geométricas em estímulos táteis, já que, na ausência da visão, o tato torna-se o mecanismo primário de percepção, seguido da audição como mecanismo secundário.

Klatzky et al. (1985) demonstram que o sistema háptico humano é bastante eficiente para a identificação de objetos. Foi realizado um teste com 20 usuários e selecionados 100 objetos de fácil identificação visual e cujos nomes são amplamente conhecidos e não ambíguos (por exemplo, pente, carta de baralho, meia, ovo, martelo etc.). Os usuários, todos com os olhos vendados, deveriam

dizer o nome de cada um dos objetos apenas manipulando-os com as mãos. Os resultados indicam que 96% das respostas foram corretas, 68% das respostas corretas foram dadas em menos de 3 segundos de manipulação e apenas 6% demoraram mais de 5 segundos. Esta experiência mostra que o tato humano é bastante rápido e preciso para identificar objetos familiares do mundo real. Entretanto, isto não significa que a mesma eficiência pode ser obtida em outras situações. Os autores mostram que o incluir um objeto em uma determinada categoria é muito diferente de obter informações estruturais sobre mesmo.

Lederman & Campbell (1982) mostram experimentos realizados com deficientes visuais para medir a utilidade de gráficos de linha em alto relevo para a representação de dados científicos bidimensionais (Figura 2.3). Em contraste com a pesquisa anterior, este trabalho investiga a capacidade do tato humano (embora voltado para o caso particular de deficientes visuais) inferir propriedades estruturais de objetos, tais como curvatura, inclinação, posição etc.

Os resultados demonstram que representações em alto relevo são úteis para deficientes visuais, sem no entanto afirmar que tais representações são suficientes para a construção de um modelo mental da estrutura espacial dos objetos como um todo. Informações locais sobre a configuração de curvas, tais como inclinação, convergência, paralelismo etc., podem ser obtidas com razoável precisão através do tato, embora em tempos bem maiores (quando comparados com a identificação de objetos familiares).

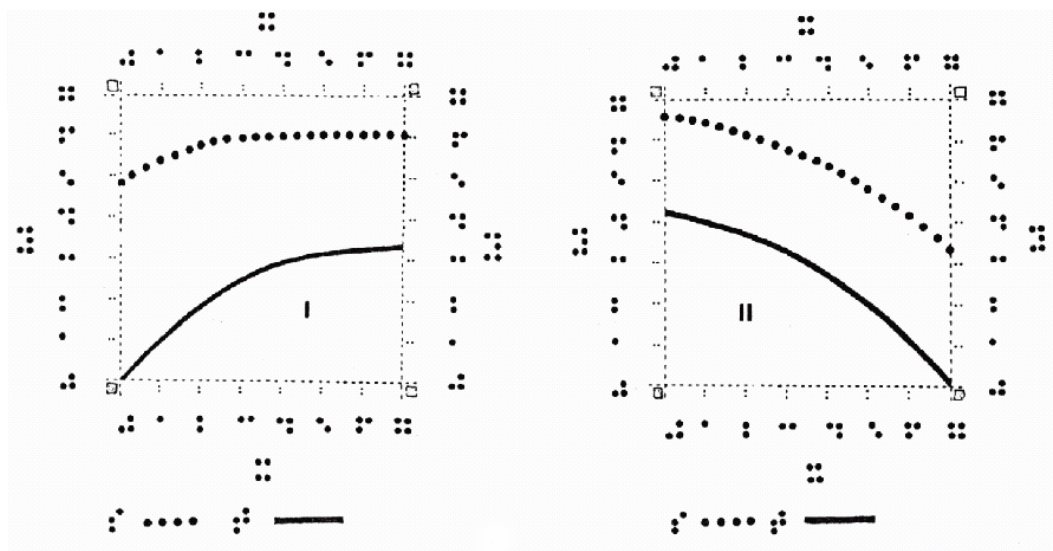


Figura 2.3: Gráficos em alto relevo (Lederman & Campbell, 1982)

2.4 Procedimentos Exploratórios

Os métodos de representação háptica citados anteriormente devem ser desenvolvidos baseados nas características básicas do tato humano. Portanto, é necessário entender como ocorre a percepção de propriedades de objetos do mundo real através do tato. Sabe-se que a percepção tátil ocorre através de sensores localizados na superfície da pele, tendões, músculos e juntas. A diversidade de tais sensores permite ao ser humano identificar vibrações, distribuição espacial da pressão em diversos pontos da pele em diferentes resoluções, distribuição do fluxo de temperatura e ainda explorar o mundo através da movimentação dos braços, dedos etc e aplicação de forças (Klatzky & Lederman, 2000).

Este rico conjunto de mecanismos sensoriais permite ao ser humano perceber diversas propriedades de objetos, verificação e monitoramento de atividades em andamento, construção de modelos mentais na ausência de condições de visibilidade etc. (MacLean, 2000). Uma das informações mais ricas proporcionadas sistema háptico humano é, sem dúvida, a percepção de propriedades físicas de objetos do mundo real, tais como textura, rigidez, temperatura, forma etc.

O estudo realizado por Lederman & Klatzky (1987) mostra que existem padrões bem típicos de movimento de mão e dedos que as pessoas intuitivamente utilizam para perceber as mais diversas propriedades físicas identificadas pelo tato. Este trabalho propõe uma taxonomia para os movimentos de mão, chamados procedimentos exploratórios (*Exploratory Procedures* ou EPs). A classificação obtida define os seguintes EPs:

- *Lateral motion*. Movimento lateral das mãos, geralmente rápidos e no interior da superfície do objeto. Está associado com a percepção de texturas (incluindo também outras propriedades relacionadas, tais como aspereza);
- *Pressure*. Aplicação de força, geralmente com os dedos, sobre a superfície do objeto. Está associado com a percepção de rigidez;
- *Static contact*. Repouso da mão sobre a superfície do objeto, sem envolvê-lo. Está relacionado com a percepção da temperatura;
- *Unsupported holding*. O objeto é levantado e mantido na mão, sem envolvê-lo. Está relacionado com a percepção do peso;

- *Enclosure*. Toda a superfície da mão toca o objeto, com a intenção de envolvê-lo. Está relacionado com a percepção de volume e forma global;
- *Contour following*. A mão mantém contato com as arestas do objeto, com a intenção de percorrê-las por inteiro. Está relacionado com a sensação da forma exata (Figura 2.4).

Os padrões apresentados acima são utilizados de maneira muito semelhante por pessoas cegas e videntes. Em condições adversas de visibilidade (por exemplo, durante a operação de uma câmera escura, durante um mergulho no rio ou ao tentar encontrar objetos no fundo de uma gaveta) pessoas videntes certamente já experimentaram os movimentos acima.

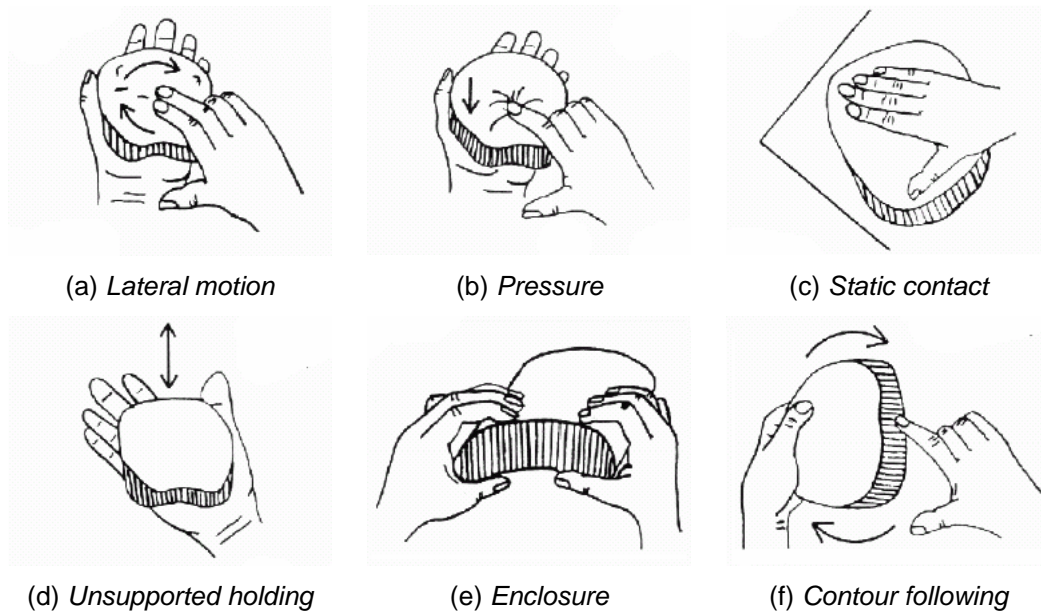


Figura 2.4: *Exploratory Procedures* (Lederman & Klatzky, 1987)

O'Modhain (1999) complementa este trabalho mostrando o que ocorre, em termos de eficiência, quando a exploração háptica de objetos do mundo real torna-se mais limitada. Os testes com um grupo de 25 usuários consistiam em classificar oito cubos do mundo real segundo as propriedades de tamanho, rigidez e textura. As condições impostas para a exploração foram divididas em cinco categorias: exploração utilizando duas mãos; uma mão; dois dedos (polegar e indicador); um dedo (indicador); uma sonda (*probe*) (Figura 2.5).

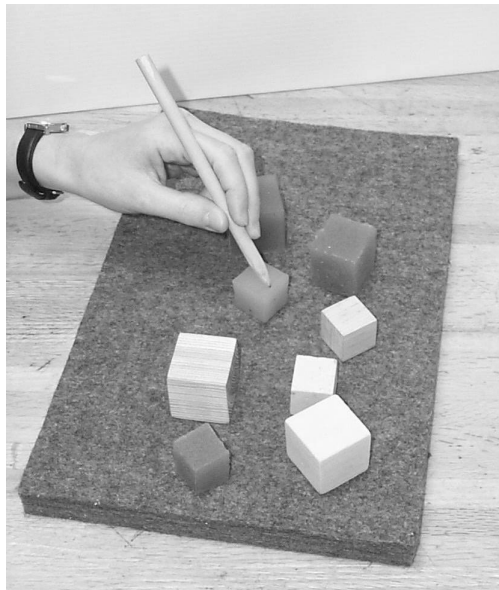


Figura 2.5: Exploração utilizando uma sonda (O’Modhrain, 1999)

Esses testes têm a finalidade de verificar a eficiência dos procedimentos exploratórios identificados por Lederman & Klatzky (1987) em condições restritas de exploração, tipicamente impostas por tecnologias hápticas (interfaces e dispositivos) utilizadas em ambientes virtuais. Tais tecnologias privam os usuários de mecanismos de *feedback* essenciais, pois muitas propriedades físicas importantes são perdidas ou mal representadas. O caso mais restrito de exploração (utilizando uma sonda) é próximo ao empregado em aplicações utilizando dispositivos hápticos 3D (por exemplo, PHANToM), conforme será visto mais adiante no Capítulo 3 desta tese. Neste caso, a sonda inviabiliza a utilização do procedimento exploratório *enclosure* (Figura 2.4e) e limita a percepção háptica em um único ponto por vez.

O’Modhrain sustenta a idéia de que os *affordances* oferecidos por um determinado dispositivo afetam os procedimentos exploratórios utilizados pelo usuário e ainda influenciam a percepção de propriedades físicas dos objetos manipulados. Além disso, mostra que quando um determinado procedimento exploratório não está disponível, os usuários tendem a criar procedimentos alternativos, ao invés de naturalmente utilizar o melhor procedimento exploratório disponível (não-otimizado, porém suficiente) conforme o esperado por Lederman & Klatzky (1987).

O entendimento dos procedimentos exploratórios apresentados acima será útil para a especificação de uma interface assistida para deficientes visuais,

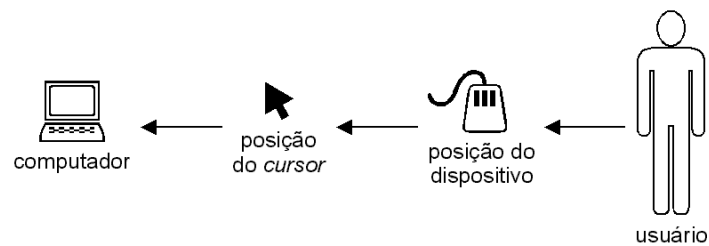
conforme será apresentado no Capítulo 6. Por isso, atenção especial será dada aos procedimentos exploratórios utilizados para a percepção de contorno (*Contour following*, Figura 2.4f), textura (*Lateral motion*, Figura 2.4a) e rigidez (*Pressure*, Figura 2.4b).

2.5 Interfaces Hápticas

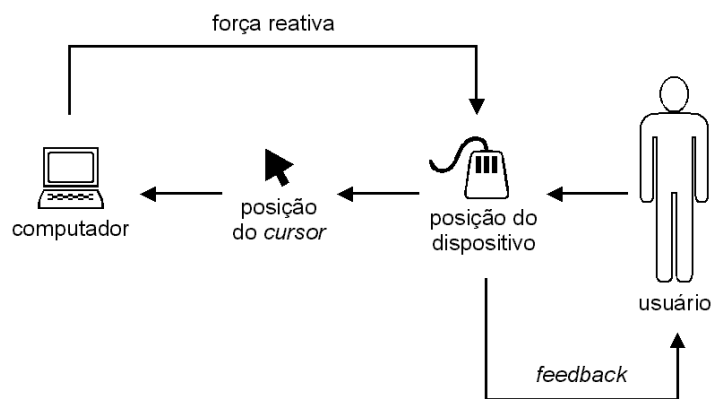
Feedbacks visuais e auditivos somente não são suficientes para permitir um usuário interagir com o computador de uma forma natural, semelhante ao processo de interação no mundo real. Isto acontece porque uma parcela significativa de nossa capacidade de estabelecer modelos cognitivos de objetos no mundo real está intrinsecamente relacionada com o sentido do tato (Massie, 1998a).

O tato é o único dos sentidos que permite *input* e *output* simultâneo, ou seja, a interação é bidirecional. Ao manipular um objeto qualquer, o usuário tanto recebe informação do mesmo (forma, peso, textura etc.) como também pode reagir de acordo com as informações recebidas (mover, rodar, deformar o objeto etc.). As interfaces convencionais utilizam interação em uma única direção (Figura 2.6a), porém uma interface reativa, que suporte o sentido do tato, pode tirar muito proveito da bidirecionalidade da interação, aumentando sensivelmente a largura de banda na transferência de informação entre o usuário e a interface (Figura 2.6b).

Os dispositivos usuais de interação, tais como *mice*, teclados e *joysticks*, atuam exclusivamente como entrada de dados. Posteriormente, com o aparecimento de sistemas de realidade virtual, surgiram *data gloves*. Estes dispositivos permitiam ao usuário interagir com o espaço tridimensional através de gestos com as mãos. Entretanto, contando apenas com o *feedback* visual, o uso de tais luvas não era eficiente, pois os usuários tinham a impressão de estarem interagindo com o vazio. Faltavam dispositivos ativos que permitissem controlar fisicamente a interação e/ou transmitir ao usuário alguma informação tátil.



(a) interação unidirecional (padrão)



(b) interação bidirecional (*force feedback*)

Figura 2.6: Tipos de Interação (Mensvoort, 2002)

Oakley et al. (2000) realizaram um experimento com pessoas de visão normal para verificar o efeito de quatro diferentes formas de *feedback* háptico em uma tarefa simples de apontamento (mover o *cursor* até a posição de um botão da interface). Os resultados mostram que o *feedback* háptico não proporciona aumento da velocidade de execução da tarefa. Entretanto, a taxa de erros diminuiu significativamente.

Fraser & Gutwin (2000) outro experimento interessante para medir os efeitos de cinco diferentes níveis de *feedback* (nenhum, auditivo, visual, tátil e combinação de auditivo, visual e tátil) em tarefas simples de seleção de objetos utilizando *mouse*. Os resultados mostram que pessoas de visão normal gostam de redundância de *feedback*, apesar de não ter proporcionado melhora significativa no desempenho do usuário (rapidez na execução das tarefas e redução nos índices de erro). Entretanto, quando esta mesma experiência é realizada em condições visuais muito desfavoráveis, nota-se nítida melhoria no desempenho dos usuários a medida em que se aumenta a redundância de *feedbacks*.

A vantagem de uma interface multimodal não é apenas a redundância. A utilização de vários canais sensoriais pode ser otimizada separando as

informações e enviando-as aos canais adequados, de acordo com as características da informação (ETSI, 2002). Assim, o canal visual, tipicamente o mais sobrecarregado, pode ser drasticamente aliviado. Esta característica é essencial para um usuário deficiente visual (Bellik & Burger, 1995; Smith et al., 1996).

Esses resultados demonstram claramente que a utilização de outros canais sensoriais pode aumentar bastante a eficiência de uma interface quando o usuário possui algum tipo de deficiência. No caso de usuários com visão subnormal, a solução mais comum é o uso de ampliadores de tela, ou seja, programas que aumentam uma determinada região da tela do computador, melhorando assim o desempenho do restrito canal visual de tais indivíduos (inLARGE, 2002; MAGic, 2002; ZoomText, 2002). No entanto, essa técnica é inútil quando o usuário não possui o canal visual, ou seja, é cego total. Neste caso, deve-se confiar em primeiro lugar no canal do tato e, como mecanismo complementar, no canal auditivo. Esta combinação é essencial em qualquer interface projetada especificamente para cegos.