

SISTEM BERSENSOR GANDA

1. PENDAHULUAN

Penggunaan saluran sensor secara bersamaan akan meningkatkan bandwidth interaksi antara manusia dan komputer. Hal ini juga akan membuat interaksi manusia-komputer menjadi seperti interaksi antara manusia dan lingkungannya.

2. USABLE SENSORY INPUTS

Ada 5 yaitu:

- Berkenaan dengan penglihatan (*visual*)
- Suara (*sound*)
- Pendengaran (*hearing*)
- Rasa (*taste*)
- Bau (*smell*)

Visual merupakan sensor yang paling terutama digunakan dalam komunikasi

Sound digunakan terbatas pada beberapa interface. Misal, suara "*beep*" digunakan sebagai "*warning*"

Taste dan *smell* merupakan sensor yang paling sedikit digunakan. Kedua sensor ini lebih digunakan sebagai penerima informasi dari pada untuk komunikasi.

Karena hanya ada sedikit cara mengimplementasikan device yang dapat me-generate keduanya, maka sensor-sensor ini tidak dikembangkan.

3. MULTI-MODAL DAN MULTI-MEDIA SYSTEM

Multi-modal system dikembangkan untuk mengambil keuntungan atas indera alami manusia. Dengan menggunakan lebih dari 1 indera atau mode komunikasi, sistem ini lebih memberdayakan penggunaan *auditory channel* (jalur yang berhubungan dengan pendengaran, suara).

Multimedia system menggunakan sejumlah media yang berbeda untuk mengkomunikasikan informasi.

Yaitu:

- Sound
- Video
- Text
- Grafik
- Icon
- Animasi
- CD-i

Multimedia system sering disebut juga sebagai Multi-modal, tapi tidak selalu.

4. PERKATAAN (SPEECH) DI INTERFACE

Berguna untuk keadaan:

- Tangan si pembicara sibuk
- Diinginkan mobilitas
- Mata si pembicara harus memperhatikan sesuatu
- Kondisi yang keras atau terkungkung yang tidak memungkinkan penggunaan keyboard.

4.1. STRUKTUR PERKATAAN (STRUCTURE OF SPEECH)

Struktur dasar:

- Bahasa Inggris terdiri dari 40 phonem.
- Tiap phonem merepresentasikan suara/bunyi yang berbeda (24 konsonan dan 16 vokal)

Perubahan pada bunyi dan kualitas phonem disebut *prosody* (intonasi) dan digunakan untuk menyampaikan banyak arti dan nuansa emosi pada suatu kalimat. *Prosody* memberikan kekayaan dan tekstur bahasa, tapi sangat sulit untuk dihitung secara kuantitas,

Co-articulation adalah cara keluarnya suara akibat dari gerakan mulut, hidung dan kecepatan bicara.

Co-articulation dapat menghasilkan berbagai macam suara yang dapat digunakan untuk membentuk allophone. Allophone melambangkan semua bunyi-bunyian dalam sebuah bahasa.

4.2 PENGENALAN PERKATAAN (SPEECH RECOGNITION)

Banyak usaha untuk membangun Speech Recognition System, tapi banyak menemui ketidaksuksesan.

Penyebab ketidaksuksesan tersebut adalah:

- Kompleksitas dari bahasa.
- Sebab lain yang lebih mengarah ke sehari-hari adalah masalah yang berkaitan dengan pengenalan secara otomatis kata yang diucapkan. Suara-suara background akan dapat mempengaruhi input, yaitu akan menutupi atau mengganggu informasi. Pembicara yang tidak jelas dapat memberikan redundansi atau suara yang tidak berarti ke dalam informasi dengan cara mengulang kata-katanya, berhenti atau menggunakan suara-suara seperti “Ummm” dan “Errrr” untuk mengisi jeda pada suatu pembicaraan.

- Variasi individu. Orang mempunyai suara yang unik, dan sistem yang diarahkan untuk sensitif terhadap variasi dan frekuensi dari suara pembicara yang sudah biasa dikenal oleh sistem dapat gagal untuk mengenali pembicara yang baru.
- Aksen dan penggunaan bahasa yang berbeda dapat menyebabkan masalah yang serius pada sistem.

Menimbang Speech Recognition dari sisi Multi Modal Interaction tidak ragu bahwa ini dapat menawarkan pilihan berkomunikasi untuk menambah cara-cara yang sudah ada. Ketika tangan user sibuk, berbicara mungkin dapat menjadi media input yang ideal. Berbicara tidak membutuhkan keyboard sehingga dalam situasi yang tertentu ada potensi untuk menggunakan sistem seperti itu. Sistem ini juga memberikan alternatif bagi user yang mempunyai cacat fisik. Seperti penderita Dyslexia.

Sampai saat ini, sistem seperti ini belum begitu sukses untuk penggunaan umum, apalagi untuk dikomersilkan.

4.3. PERKATAAN BUATAN (SPEECH SYNTHESIS)

Speech synthesis merupakan pelengkap dari speech recognition. Ide agar dapat berbicara dengan komputer merupakan hal yang menarik bagi banyak user, khususnya bagi mereka yang tidak bisa komputer.

Masalah yang ada pada speech synthesis:

- User sangat sensitif terhadap variasi dan informasi bicara. Sebab itu mereka tidak dapat mentolerir ketidaksempurnaan pada speech synthesis.
- Output dalam bentuk suara (spoken output) tidak dapat diulang atau di browse dengan mudah.
- Meningkatkan noise (berisik) pada lingkungan kantor. Atau bila menggunakan headphone akan meningkatkan biaya.

Lingkungan aplikasi speech synthesis:

- Bagi tuna netra, speech synthesis menawarkan media komunikasi dimana mereka memiliki akses yang tidak terbatas.
- Lingkungan dimana visual dan *haptic skill* user sedang terfokus pada hal lain. Contohnya signal bahaya pada aircraft cockpit.

5. NON-SPEECH SOUND

Digunakan sebagai alarm dan warning, atau status information. Penampilan informasi yang redundan dapat meningkatkan kinerja user. Contohnya, user dapat mengingat suara yang mencerminkan icon tertentu, tapi bukan tampilan visualnya.

Ketidakjelasan satu mode dapat dipecahkan dengan menggunakan informasi yang ditampilkan melalui media lain. Misal: speech recognition system juga menggunakan kamera untuk merekam gerakan bibir. Kata-kata yang tidak jelas atau frase dapat dipecahkan lebih akurat dengan menggunakan informasi visual.

Perbedaan speech dan non-speech:

Speech: serial dan user harus mendengarkan keseluruhan kalimat sebelum mengerti maksudnya, sehingga membutuhkan waktu yang panjang.

Non-speech:

- Dapat diasosiasikan dengan kejadian khusus, dan waktu yang dibutuhkan lebih pendek
- Universal (mempunyai arti yang sama dibelahan bumi manapun)
- Menyediakan informasi status
- Navigasi sistem

5.1. SOUNDTRACK

Adalah suatu word-processor dengan sebuah auditory interface, yang didisain untuk user yang bermasalah dengan alat visualnya.

File Menu	Edit Menu	Sound Menu	Format Menu
Alert	Dialog	Document1	Document2

Soundtrack's Main Screen

Tiap sel memiliki tone yang berbeda jika kursor ada di sel tersebut. Dengan menggunakan tone-tone ini, user dapat menjelajahi sistem.

Dengan meng-klik suatu sel, sel tersebut akan menyuarakan nama sel-nya. Hal ini dapat memberikan informasi yang pasti yang dapat mengorientasikan kembali user yang bingung atau 'lost'.

Dengan melakukan double klik, maka akan muncul sub-menu dari item tersebut. Item-item di sebu-menu juga memiliki tone; semakin bergerak turun, tone-nya menjadi rendah, dan bila naik tone-nya menjadi tinggi.

Cara memasukkan text ke dalam soundtrack adalah dengan cara menyebutkan kata-kata atau karakternya. User menghitung tone-tone yang berbeda untuk mengetahui lokasi/posisi merek di screen.

Soundtrack menyediakan auditory solution untuk menampilkan word-prossecor secara visual, walaupun hasilnya tidak semuanya dapat divisualisasikan di interface.

5.2. AUDITORY ICONS

Auditory icon menggunakan suara alami untuk merepresentasikan tipe-tipe obyek dan aksi yang berbeda di interface.

Pada perusahaan SonicFinder, auditory icon digunakan untuk merepresentasikan desktop object dan action.

Contoh: suara benturan digunakan untuk aksi membuang sesuatu ke keranjang sampah. User mempelajari suara-suara ini dengan mengira/menebak familiar action dalam kehidupan sehari-hari.

Permasalahan auditory icon adalah beberapa objek dan action tidak memiliki kejelasan, suara alami untuk mengidentifikasikannya.

Pada perusahaan SharedARK, sound digunakan dalam 3 hal:

- Sebagai confirmation of actions
- Status information
- Aids to navigation

Suara penegasan (Confirmatory sound) menyediakan informasi yang redundan yang akan meningkatkan feedback.

Proses dan *state information sound* ada dalam 2 level, yaitu:

- Global sound : merepresentasikan state keseluruhan sistem dan dapat didengar dari mana saja.
- Local sound : spesifik untuk eksperimen khusus dan berubah saat user merubah 1 eksperimen ke eksperimen yang lain.

Navigational information diletakkan dimana saja di sistem dan akan semakin keras jika user bergerak mendekat, dan volumenya akan pelan ketika user menjauh.

5.3. EARCONS

Sebuah alternatif penggunaan suara alami adalah untuk merancang synthetic sounds. Earcons menggunakan kombinasi terstruktur dari not/nada, yang disebut *motives*, untuk merepresentasikan aksi dan obyek.

Ada 2 tipe kombinasi earcons:

- **Compound earcons** : menggabungkan motive yang berbeda untuk membuat sebuah aksi yang spesifik.
Contoh: mengkombinasikan motives "create" dan "file"
- **Family earcons** : menampilkan compound earcon yang setipe.
Contoh: operating system error dan syntax error berada dalam "error" family.

Dengan cara ini earcon dapat dibuat terstruktur secara hirarki, untuk merepresentasikan menu.

Earcon dengan mudah dapat dikelompokkan dan diperbaiki sesuai dengan komposisi dan hirarkinya. Tapi earcon susah untuk dihubungkan dengan tugas yang spesifik di dalam interface karena ada mapping yang tidak beraturan.

6. PENGENALAN TULISAN TANGAN (HANDWRITING RECOGNITION)

Seperti speech, tulisan tangan merupakan bentuk alami dari komunikasi. Tulisan tangan menawarkan input tekstual dan input grafik dengan menggunakan alat yang sama.

6.1. TEKNOLOGI

Teknologi yang digunakan adalah *digitizing tablet*. Coretan yang dibuat dengan pena akan ditransformasikan ke dalam seri koordinat, kira-kira 1 dalam 1/50 detik. Gerakan-gerakan yang cepat akan menghasilkan titik-titik yang lebih jarang dibandingkan dengan gerakan-gerakan lambat.

Digitizing tablet telah diperbaiki dengan menggabungkan sebuah screen tipis untuk menampilkan informasi, yang akan membentuk kertas elektronik (electronic paper).

Informasi yang ditulis di digitizer dapat dengan mudah di tampilkan kembali (re-display), atau disimpan dan di re-display untuk referensi lebih jauh.

System akan bermanfaat jika dapat menginterpretasikan coretan yang diberikan dan menghasilkan text.

6.2. PENGENALAN TULISAN TANGAN

Tulisan tangan orang per orang sangatlah bervariasi, dan tiap orang berubah-ubah tulisan tangannya dari hari ke hari. Hal ini akan menimbulkan kesulitan untuk sistem pengenalan, yang bekerja dengan mencoba mengidentifikasi baris-baris yang berisi teks, dan kemudian memilah-milah image yg didigitasi ke dalam karakter-karakter yang terpisah.

Walaupun demikian, ketika tulisan ditulis secara individual, dengan sedikit pemisahan, kesuksesan sistem mejadi lebih layak, walaupun sistem harus dilatih untuk mengenali karakterisitk bermacam-macam user.

Banyak solusi-solusi yang diusahakan pada *speech recognition* juga dicoba pada *handwriting recognition system*, dan begitu pula sebaliknya. Seperti, pengenalan keseluruhan kata, penggunaan konteks untuk karakter-karakter yang yang memiliki banyak arti, dan neural networks, yang dipelajari dengan contoh.

7. TEXT, HYPERTEXT DAN HYPERMEDIA

Output tekstual ditampilkan dalam bentuk sebuah blok teks, atau sebagai suatu urutan halaman yang harus dibaca. Bagaimanapun menampilkan teks dalam gaya berurut, seperti halaman per halaman, adalah tidak cocok untuk browsing, karena hal ini memaksa pembaca untuk mengikuti alur berpikir penulis. Hal ini menimbulkan kesulitan. Misalnya ketika pembaca sedang membaca buku manual. Jika pembaca tersebut tidak mengerti akan satu istilah dalam halam itu, ia harus mencari keterangan atas istilah tersebut dalam halaman-halaman yang lain. Setelah itu dia baru bisa kembali ke halaman yang tadi dan melanjutkan kegiatan membacanya.

Hypertext mencoba membatasi keterbatasan-keterbatasan pada text dengan cara membentuk ke dalam web screen dari pada secara baris. Cara ini membolehkan halaman-halaman yang berbeda diakses dari halaman yang sedang aktif, dan jika didisain dengan baik, user akan merasa lebih mudah untuk mengikuti suatu pokok bahasan yang diinginya.

Seperti text, hypertext juga mengandung diagram, photo dan media lainnya. Sistem seperti ini dikenal sebagai *hypermedia system*.

Hypertext terdiri dari beberapa halaman, dan satu set links yang digunakan untuk menghubungkan halaman-halaman menjadi satu.

Hal yang perlu digaris bawahi adalah bahwa user dapat membaca dokumen dengan cara yang dirasa paling baik untuknya.

Hypertext dan hypermedia adalah satu jaringan simpul (disebut artikel, dokumen, file, kartu, halaman, frame, layar) yang berisi informasi (dalam bentuk teks, grafik, video, sound, dll) yang dihubungkan dengan link (disebut juga cross-reference, citation)

Hypertext umumnya diterapkan untuk aplikasi teks.

Hypermedia umumnya digunakan untuk sistem yang meliputi media lainnya, terutama sound dan video.

7.1. APLIKASI HYPERMEDIA

- **Hypercard**
Yang meniru sebuah kartu indeks, yang dari situ user dapat mencari sesuatu. Masing-masing kartu dapat menyimpan text, diagram, photograph, bitmap dan lain-lain. Hot spot pada kartu ini mengijinkan gerakan antar kartu. Selain itu, pada kartu-kartu ini juga terdapat tombol forward dan backward dan sebuah icon "home", untuk membolehkan user bergerak secara sequensial
- **Information management**
- **Sistem pendidikan**
- **Sistem aplikasi pembantu (Help System Application)**

7.2. MASALAH PADA HYPERMEDIA

- **Navigating the hypertext web.**
Sangatlah sulit untuk memutuskan dengan pasti posisi seseorang di dalam hypertext web. Solusi untuk masalah ini adalah menyediakan map dari dokument hypertext yang mengidentifikasikan posisi seseorang. Links ke awal atau akhir dapat diidentifikasikan dan user akan sedikit sekali akan "lost".
- **Jika informasi disajikan dalam struktur yang lebih kompleks, terdapat kemungkinan terlewatkannya halaman-halaman tertentu atau item-item suatu informasi.**

Kedua masalah ini telah dipecahkan dengan menyediakan rute yang diusulkan untuk dibaca user melalui hypertext dokument. Tapi hal ini justru berlawanan dengan tujuan dibuatnya sistem seperti ini.

8. ANIMASI DAN VIDEO

8.1. ANIMASI

Animasi menjadi elemen yang paling biasa / lazim pada sebuah sistem. Dengan animasi dapat membuat suatu image berpindah, berubah (alter) dan berganti pada waktu yang bersamaan.

Contoh:

- **Animasi kursor**
- **Jam**
- **Icon jam gelas untuk menunjukkan suatu pekerjaan sedang berlangsung**
- **Kartun**

8.2. VIDEO DAN DIGITAL VIDEO

Penggunaan video akan memperluas potensi interface dan memberikan designer tool/alat lain, dan cara lain untuk menampilkan informasi.

Macam-macam CD:

- **CD-I**
 - Mempunyai video yang terbatas dan kapasitas still image
 - Ditujukan untuk pasaran domestik
 - Baik untuk full screen animation dari pada untuk video work.
- **CD-XA (Extended Architecture)**
 - Pengembangan dari CD-I
 - Lebih mendukung audio digital dan still image
- **DVI (Digital Video Interactive)**
 - Menampilkan full motion video, bersaing dengan UVC (Universal Video Communications)
 - Dapat menempatkan full motion video ke dalam single CD lebih dari 1 jam, dan teknologinya meningkatkan kompresi dan kapasitas penyimpanan sepanjang waktu.

Contoh dari penggabungan video ke dalam interface adalah QuickTime dari Apple.

9. PENGENALAN GERAK (GESTURE RECOGNITION)

Gerakan adalah komponen IMK yang menjadi pusat perhatian dalam Multi-modal system. Dengan beberapa gerakan untuk mengontrol komputer, adalah merupakan suatu keuntungan pada situasi dimana tidak ada kemungkinan untuk melakukan pengetikan atau ketika panca indera lainnya tidak bisa digunakan.

Beda dengan *speech* dan *handwriting*, gerakan merupakan:

- User-dependent
- Subject to variation
- Co-articulation

Teknologi yang digunakan untuk menangkap suatu gerakan adalah:

- Vision system
- Dataglove

Kedua teknologi ini sangat mahal dan sulit untuk menginterpretasikan data.

Sebuah team dari Toronto membuat sebuah sistem pengenalan gerakan yang men-translate gerakan tangan ke dalam *synthesized speech*, dengan menggunakan 5 (lima) jaringan syaraf yang bekerja secara paralel. Kelima jaringan syaraf ini mempelajari dan mengartikan bagian-bagian input yang berbeda.

10. COMPUTER VISION

- Dengan vision system yang baik, memungkinkan komputer untuk mengenali user-nya dan membentuk sistem yang sesuai dengan permintaan. Computer vision mempunyai efek yang dalam pada interaksi.
- Computer vision system cenderung berdasar pada pendekatan “bottom-up” dimana image-image diperlakukan sebagai satu set pixel yang secara progresif dianalisa untuk mengambil inti dari artinya.
- Satu set tertentu dari proses akan dapat mengambil input image dan memprosesnya untuk menghilangkan noise.

11. APLIKASI MULTI-MEDIA SYSTEM

11.1. INTERFACE UNTUK USER DENGAN KEBUTUHAN KHUSUS

- Penelitian interface untuk user dengan kebutuhan-kebutuhan khusus terfokus pada penyediaan sistem interaksi bagi mereka yang kebutuhannya tidak disediakan oleh komputer standard. Biasanya hal ini karena interface menggunakan saluran input atau output yang tidak dapat digunakan dengan efektif oleh user.
- Berkembangnya penggunaan graphical interface mengurangi kemungkinan bagi user yang mengalami kerusakan visual/penglihatan.
- Standard interface saat ini adalah graphical interface. Karena interface yang berkembang saat ini banyak menggunakan gambar-gambar/grafik, maka akses ke komputer bagi user yang mengalami kesulitan penglihatan malah berkurang. Hal ini disebabkan output dalam bentuk Braille tidak dapat digunakan untuk merepresentasikan gambar. Beberapa sistem mencoba untuk menyediakan akses ke graphical interface untuk kelompok user ini, dengan cara menambahkan suara ke interface. Sistem seperti ini telah sukses, tapi memiliki kekurangan ketika menghadapi banyak data yang harus direpresentasikan.
- Untuk user yang mengalami gangguan bicara dan pendengaran, multi-media sistem menyediakan beberapa alat komunikasi, termasuk synthetic speech dan komunikasi berdasarkan teks dan conferencing system.

- User dengan bermacam gangguan fisik berbeda-beda kemampuannya untuk mengontrol dan bergerak, tapi kebanyakan dari mereka susah dalam menggunakan mouse.
- Speech input dan output adalah salah satu pilihan bagi yang tidak mengalami gangguan bicara. Alternatif lain adalah sistem Eyegaze yang mengikuti gerakan mata untuk mengontrol kursor, atau driver keyboard yang dipasang di kepala user. Jika user tidak dapat mengontrol gerakan kepala, gerakan isyarat dan movement tracking dapat digunakan oleh user sebagai alat kontrol. Jika user mempunyai keterbatasan menggunakan keyboard, sebuah sistem yang dapat memprediksi (predictive system) seperti Reactive keyboard, dapat menolong, dengan cara mengantisipasi command-command yang akan diketik dan menawarkan perintah itu untuk dieksekusi. Ini dapat meningkatkan pengetikan kata. Perkiraan itu didasarkan atas apa yang telah user ketik sebelumnya, seperti dalam Excel.
- User yang mempunyai kesulitan belajar, seperti dyslexia, akan menemui kesulitan menghadapi informasi dalam bentuk teks. Pada kasus yang parah, speech input dan output dapat mengurangi kebutuhan untuk membaca dan menulis, dan meningkatkan akurasi input dan output. Pada kasus yang tidak begitu berat, fasilitas koreksi ejaan dapat menolong user.

11.2. VIRTUAL REALITY

- Virtual reality berhubungan dengan simulasi yang diciptakan komputer atas suatu situasi dimana user seolah-olah berada dalam situasi itu. VR ini merupakan seni dalam Multi-Media System tapi lebih ditekankan pada sensasi visual.
- Sebagai input, digunakan sarung tangan yang dapat memberikan informasi berdasarkan gerakan. Umpan balik dapat disertakan ke sarung tangan tersebut, sehingga seolah-olah user dapat merasakan hentakan ketika menggenggam sesuatu objek. speech recognition system dapat juga diikutsertakan dan secara umum VR-system dilengkapi oleh audio feedback.

Salah satu aplikasi dari VR adalah:

- Permainan 3 dimensi
- Data visualization. Umumnya digunakan pada bidang protein chemistry
- Simulasi penerbangan (merupakan aplikasi VR secara interaktif)