

República Federativa do Brasil
Ministério do Desenvolvimento, Indústria
e do Comércio Exterior
Instituto Nacional da Propriedade Industrial.

(21) **PI0805654-4 A2**



* B R P I 0 8 0 5 6 5 4 A 2 *

(22) Data de Depósito: 30/12/2008
(43) Data da Publicação: 14/09/2010
(RPI 2071)

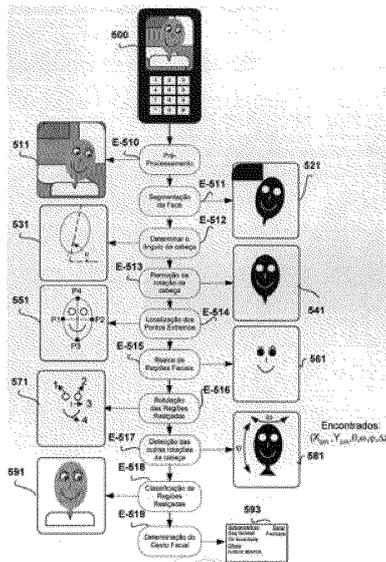
(51) *Int.Cl.:*
G06K 9/00

(54) **Título: MÉTODO DE RASTREAMENTO FACIAL E IDENTIFICAÇÃO DE FEIÇÕES EM DISPOSITIVOS PORTÁTEIS**

(73) Titular(es): Samsung Eletrônica da Amazônia Ltda.

(72) Inventor(es): Cincinato Furtado Leite Neto, César Lincoln Cavalcante Mattos, Fabio Cisne Ribeiro, Igor Rafael Silva Valente, José Marques Soares, Kim Lima Silva, Paulo Cesar Cortez, Robson da Silva Siqueira, Rodrigo Carvalho Souza Costa

(57) **Resumo:** MÉTODO DE RASTREAMENTO FACIAL E IDENTIFICAÇÃO DE FEIÇÕES EM DISPOSITIVOS PORTÁTEIS. A presente invenção diz respeito a um método de detecção e rastreamento da face e gestos faciais a partir de técnicas de segmentação de imagens. O método da presente invenção visa fornecer uma forma prática para interagir com dispositivos portáteis através da utilização de movimentos da cabeça. O método envolve a captura de uma seqüência de imagens, a segmentação da face existente nas imagens, o rastreamento dos movimentos realizados pela cabeça e reconhecimento de expressões faciais. O método proposto é capaz de controlar aplicações no próprio dispositivo portátil, simulando um controle virtual, em que os comandos são executados de acordo com os gestos e feições faciais reconhecidas. Isto permite o controle da navegação de menus e jogos, dentre outras aplicações.





Relatório Descritivo da Patente de Invenção para: **"MÉTODO DE RASTREAMENTO FACIAL E IDENTIFICAÇÃO DE FEIÇÕES EM DISPOSITIVOS PORTÁTEIS"**.

Campo da Invenção

5 A presente invenção está relacionada ao campo da interação humana com dispositivos portáteis, mais especificamente à interação feita através do reconhecimento de gestos e feições faciais realizadas por um usuário em dispositivos portáteis.

10 Antecedentes da invenção

Devido ao avanço nas tecnologias de comunicação móvel e a disponibilidade de acesso à banda larga em dispositivos portáteis, tem-se observado o aumento da procura por aplicações que utilizem essas tecnologias, destacando-se o
15 *download* de fotografias digitais, de músicas e de vídeos, o envio e recebimento de mensagens multimídia (MMS), conexão a *peer to peer*, além da navegação na *internet*, entre outras.

No caso específico do envio de uma mensagem multimídia
20 (MMS) através de um celular ou *smartphone*, o usuário grava sua imagem e envia para outro usuário. Para isto, o vídeo é comprimido utilizando a codificação em MPEG4 ou H.263. Já na vídeo-chamada, há o envio/recebimento em tempo real de vídeos adquiridos pelos dois usuários que estão se

comunicando. Para realizar esta comunicação, o usuário pode contratar uma quantidade pré-definida de tráfego de dados junto a uma operadora de telefônica celular, ou optar por pagar, de forma exclusiva, pelo número bytes de dados que venha a usar. Entretanto, atualmente, o custo do byte trafegado por uma rede de dados móvel ainda é consideravelmente elevado. No caso do tráfego de vídeos, mesmo que tenham sido previamente comprimidos, a quantidade de bytes transferidos é da ordem de milhões, tornando a utilização deste tipo de serviço muito dispendiosa para o usuário final.

Uma forma de diminuir custos e popularizar esta comunicação é a utilização de personagens animados. Eles podem representar uma alternativa ao vídeo do usuário, pois necessitam de pelo menos 50% menos transferência de dados para efetuar a comunicação. Além disso, permitem a comunicação não-verbal e, conseqüentemente, o suporte perceptivo em forma de expressões faciais durante comunicações remotas. Além disto, possibilitam que o usuário personalize a imagem de seu modelo virtual, como o tipo e cor de cabelo, formato dos olhos, forma da cabeça, etc.

No contexto da presente invenção, avatar é a representação gráfica de um usuário em realidade virtual,

que pode variar desde um sofisticado modelo tridimensional até uma simples imagem. Humanóides virtuais ou *avatars*, são personagens animados por computador que podem ser representados em duas ou três dimensões e controlados por diferentes técnicas. Além disso, pode-se animar os avatares
5 através de comandos presentes em interfaces gráficas, nas quais o usuário deve escolher os comandos a partir de um conjunto finito de botões ou através de um dispositivo de entrada de dados, como um mouse ou teclado.

10 Outra forma de animar o humanóide virtual consiste em interpretar gestos realizados pela face de um usuário, como por exemplo, balançar ou inclinar a cabeça para esquerda e direita, piscar os olhos, abrir e fechar a boca, levantar as sobrancelhas, através da aquisição e do processamento de vídeo quadro a quadro para extrair as informações de
15 movimento. E com isto, estimar o gesto realizado.

Um humanóide virtual é dividido nas seguintes partes: cabeça, tronco e membros. Para o contexto da presente invenção, apenas a cabeça será considerada.

20 A estimação da posição da cabeça é um problema tridimensional, em que os parâmetros estudados estão relacionados com as três formas de girar a cabeça. O problema da estimação da pose da cabeça tem sido estudado nas últimas décadas, sendo tratado com base em

classificadores e em características geométricas.

São conhecidos no estado da técnica métodos baseados em classificadores, como, por exemplo, no documento "*Multi-view face and eye detection using discriminant features*", de Peng Wang e Qiang JiIan, publicado na revista *Computer Vision and Image Understanding* vol. 105 em 2007. Este método utiliza a análise discriminante não recursiva e não paramétrica (RNDA) para extrair as características faciais, bem como um algoritmo de combinação de múltiplos classificadores. No entanto, os métodos baseados em classificadores indicam apenas a direção da cabeça (esquerda, frontal e direita), não possibilitando o reconhecimento de gestos faciais, na medida em que não são capazes de determinar o quanto a direção da cabeça foi modificada. Esta informação é particularmente importante para a animação do avatar, pois torna o seu movimento mais suave.

O documento de patente norte-americano US 7,266,225, de 04 de setembro de 2007, utiliza a localização do eixo do nariz do usuário para estimar a direção da face. Este método propõe a identificação do eixo que maximiza a correlação entre o lado direito e esquerdo da face. A partir do referido eixo, é detectada a direção da cabeça (cima, baixo, esquerda e direita). De acordo com os

ensinamentos do referido documento, o método de detecção da face é baseado na diferença da intensidade de níveis de cinza da imagem. Por outro lado, a presente invenção propõe a detecção da face a partir de um realce de cor através de uma relação entre os canais de cor Cb e Cr que compõe a imagem. Além disso, os pontos de referência utilizados para determinar a direção da cabeça são as narinas e não um eixo de maximização, como apresentado no documento US 7,266,225.

O documento de patente US 6,697,072, de 24 de fevereiro de 2004, utiliza um método baseado em subtração de fundo para controlar um avatar. De acordo com os ensinamentos do documento US 6,697,072, é necessário realizar uma etapa de calibração, apresentando o fundo (sem a presença do usuário) e a partir daí, o método realiza o rastreamento da pessoa. Este método não é adequado para utilização em dispositivos portáteis devido às modificações constantes no cenário em que o usuário está inserido, seja pelo seu deslocamento ou por ser difícil manter o dispositivo totalmente parado. Este problema dificulta a utilização deste método para dispositivos portáteis em geral.

O documento de patente US 6,677,969, de 13 de janeiro de 2004, descreve um método de reconhecimento gestual para animação de um humanóide virtual baseado no reconhecimento

da direção da face associado a um conjunto de botões. O reconhecimento da posição da face anima a posição da cabeça do avatar enquanto o conjunto de botões escolhe o tipo de feição. Por exemplo, um botão ativa uma feição sorridente, enquanto outro deixa a feição triste. Este método não faz o reconhecimento da feição, mas aguarda o usuário apertar o botão. Desta forma, para fazer várias feições, é necessário apertar uma série de botões, tornando a utilização deste método exaustiva.

10 O documento de patente brasileiro PI 9909611-0, de 12 de abril de 1999, descreve um método de reconhecimento de feições para animação de um avatar baseado em *Wavelets*. De acordo com o referido documento brasileiro, utilizam-se uma série de *Wavelets* para detectar pontos importantes da face e a partir destes realiza-se o rastreamento. Cada tipo de ponto da face é encontrado a partir da aplicação de uma transformada *Wavelet* específica. A implementação de transformadas *wavelets*, de modo geral, demandam um grande esforço computacional. Assim, para o seu uso em aplicações que precisam de uma resposta em tempo real, este tipo de técnica só pode ser utilizada em computadores pessoais que, em geral, possuem um poder de processamento dez vezes superior ao disponível em um dispositivo portátil, como é o caso de um telefone celular.

O documento "3D Face Pose estimation and tracking from monocular camera", proposto por Quiang Ji e Rong Hu na revista *Image and Vision Computing*, em fevereiro de 2002, utiliza a iluminação próxima do infravermelho para realizar um efeito de brilho nas pupilas de uma pessoa. A localização dos pontos brilhantes correspondentes é então utilizada para estimar a pose da cabeça do indivíduo. Por outro lado, no método da presente invenção, não é necessário a utilização de nenhum tipo de câmera especial, apenas aquela que já vem integrada ao dispositivo portátil.

Outra maneira de realizar o rastreamento facial é fazer a segmentação da cabeça e, a partir do resultado obtido, realizar o reconhecimento dos olhos e da boca. Um método de rastreamento da cabeça foi proposto por T. R. Hansen no trabalho "*Mixed Interaction Spaces - a new interaction technique for mobile devices*", apresentado no UbiComp 2005. Este trabalho utiliza um pacote de funções denominadas de "CamShift", produzido pela Intel, para detecção e rastreamento do movimento da cabeça do usuário em dispositivos portáteis.

O CamShift foi inicialmente proposto por Gary R. Bradski no trabalho "*Computer Vision Face Tracking For Use in a Perceptual User Interface*", publicado no *Intel Technology Journal* em 1998. Este método consiste em um

algoritmo interativo para estimar parâmetros de segmentação com base na distribuição de cor de uma região da imagem escolhida pelo usuário. Para que este método funcione bem, é necessário marcar uma região e escolher o tamanho da janela de busca inicial para, em seguida, realizar o rastreamento da face. Na prática, tal necessidade implica em restrições do seu uso por pessoas não treinadas e que não entendem como escolher corretamente os parâmetros de rastreamento.

Os métodos descritos anteriormente possuem a capacidade de detectar a posição (X, Y) , a distância da câmera (Z) e a rotação (θ) . No entanto, estes não realizam o reconhecimento das expressões faciais realizadas. Isto porque, a localização dos olhos, nariz e boca são dependentes da direção da cabeça e estes métodos não identificam este tipo de movimento.

Assim, de acordo com os ensinamentos da presente invenção, é proposto um método para reconhecer não somente a posição e a rotação da cabeça, mas também a detecção de feições independentemente da inclinação da cabeça sem a utilização de novos sensores eletrônicos no dispositivo.

Dessa forma, as demais operações para identificação de feições podem ser feitas considerando-se uma face com inclinação nula, o que reduz a complexidade da localização

das características faciais. Para encontrar estas características, utiliza-se um realce por pigmentações e a geometria da face. A invariância à inclinação da cabeça possibilita ainda a localização rápida dos pontos extremos da face (superior, inferior, mais à esquerda e mais à direita). Esses pontos, assim como as distâncias horizontal e vertical da face são fundamentais para limitar as regiões de busca das características faciais, o que diminui a carga de processamento do método.

10 Além disto, o método proposto possibilita ainda a aplicação do método para rastreamento de objetos. O ângulo e a localização dos pontos extremos podem ser usados para rastrear um objeto, mudando apenas os parâmetros de segmentação.

15 A aplicação do método proposto é feita em tempo real em sistemas de dispositivos portáteis com baixo poder de processamento. O método proposto não utiliza cálculos complexos como aqueles realizados para computar as transformadas de *wavelet*, possuindo assim um baixo custo de execução.

20 O método proposto pela presente invenção possibilita ainda a utilização da referida tecnologia nos dispositivos portáteis conhecidos sem a necessidade de uma modificação na estrutura do hardware, sendo necessária apenas a

atualização de software para a utilização da tecnologia.

Sumário da Invenção

A presente invenção apresenta um método de detecção e rastreamento da face e gestos faciais a partir de técnicas de segmentação de imagens. O método proposto pela presente invenção se diferencia dos métodos do estado da técnica por introduzir um método eficiente de detecção de faces e reconhecimento de movimento faciais que pode ser implementado por técnicas computacionalmente de baixa complexidade. Isto implica em um trabalho de otimização e ajuste de técnicas de processamento digital de imagens e vídeo para viabilizar sua implementação em dispositivos portáteis.

O método da presente invenção visa fornecer uma forma prática para interagir com dispositivos portáteis através da utilização de movimentos da cabeça. O método envolve a captura de uma seqüência de imagens, a segmentação da face existente nas imagens, o rastreamento dos movimentos realizados pela cabeça e reconhecimento de expressões faciais. O método proposto é capaz de controlar aplicações no próprio dispositivo portátil, simulando um controle virtual, em que os comandos são executados de acordo com os gestos e feições faciais reconhecidas. Isto permite o controle da navegação de menus e jogos, dentre outras

aplicações.

As feições faciais e movimentos realizados pela cabeça, detectados a partir da utilização do método da presente invenção, fornecerão uma alternativa para a
5 comunicação por videoconferência. Neste caso, ao invés de enviar a imagem do usuário adquirido pela câmera, apenas parâmetros para a animação de um humanóide virtual seriam enviados. Esta comunicação possui a vantagem de utilizar menos banda do que os métodos de videoconferência atuais.

10 Desta maneira, o método aqui proposto aumenta a possibilidade de formas de interação. O método proposto possibilita ainda uma nova forma de interação com os dispositivos portáteis sem a necessidade de modificação no seu projeto, seja na sua placa de montagem ou pela adição
15 de teclas extras no seu *layout*.

Além disso, a presente invenção provê um método para rastrear, em tempo real, a cabeça do usuário e promover o reconhecimento de gestos realizados pela face através de técnicas de visão artificial.

20 Assim, a presente invenção apresenta as seguintes vantagens diante das soluções conhecidas do estado da técnica:

➤ Baixo custo computacional. Em função das otimizações feitas em cada uma das etapas do método

proposto, a presente invenção pode ser executada em um dispositivo portátil sem degradar o desempenho das outras atividades que são também realizadas pelo dispositivo de forma concomitante.

5 ➤ Localização do ponto extremo da forma. As localizações das regiões são baseadas apenas nos pontos extremos, que são calculados a partir dos eixos mediais que passam pelo centro de massa. Estes pontos calculados servem para localizar as prováveis regiões em que se localizam os
10 elementos importantes da face: boca, narinas e olhos.

 ➤ Invariância à rotação e à escala. A detecção das características faciais (narinas, olhos, boca) é extraída a partir da face após a remoção da rotação, através da projeção dos pixels em relação ao eixo médio da face.

15 ➤ Detecção de vários eixos de movimento independentes. O método aqui proposto possibilita o controle de um cursor virtual no dispositivo portátil baseado na detecção da posição da cabeça e do ângulo em que ela se encontra. Movimentos nos eixos horizontal (eixo X) e
20 vertical (eixo Y) podem ser baseados na posição do centro de massa da cabeça e do ângulo de inclinação mais para direita, esquerda, olhando para cima ou para baixo. O movimento sobre o eixo ortogonal ao plano XY pode ser detectado pela variação de área da cabeça, podendo simular

um movimento de aproximação (*zoom in* e *zoom out*).

➤ Desenvolvimento de um método de interação baseado em cor. Parte do método de estimação de posição pode ser utilizado como sistema de interação por visão artificial.

5 Uma forma é utilizar o rastreamento do ângulo θ , a variação de área (ΔA) e a posição do centro de massa (X_{cm}, Y_{cm}). Assim, através de parte do método proposto, pode-se formar um sistema de interação apenas escolhendo parâmetros diferentes para segmentação por cor.

10 ➤ Realce de características faciais. Através de uma combinação dos canais C_b e C_r , são realçadas as características faciais, facilitando a localização das mesmas.

Breve descrição das figuras

15 Os objetivos e as vantagens da presente invenção tornar-se-ão mais evidentes a partir da descrição detalhada a seguir de um exemplo de concretização da invenção e desenhos anexos fornecidos a título de exemplo não-limitativo, em que:

20 A figura 1 ilustra os tipos de movimentos realizados pela cabeça.

A figura 2 é mostra um esquema de movimento angular do dispositivo portátil para geração de ações.

A figura 3 mostra um modelo genérico de dispositivo portátil, no qual a presente invenção pode ser implementada.

5 A figura 4 é um diagrama de blocos do método para reconhecimento de gestos faciais.

A figura 5 mostra uma visão geral para funcionamento de um sistema a partir da utilização do método proposto pela presente invenção.

10 A figura 6 mostra um método para detecção de pontos extremos da face.

A figura 7 mostra a região em que é aplicado o realce de características faciais.

A figura 8 mostra regiões utilizadas para localizar os olhos e sobrancelhas.

15 A figura 9 mostra pontos característicos da sobrancelha.

Descrição das concretizações preferidas da invenção

20 A figura 1 ilustra os eixos (x, y, z) e uma seta indicando em que eixo está ocorrendo o movimento de rotação. Em 101, é mostrada a face de um usuário. O primeiro tipo de movimento consiste no movimento de rotação da cabeça em relação ao eixo x (normal à face). Este movimento ocorre quando o usuário encontra-se de frente à câmera e inclina a cabeça. Este movimento é conhecido como

ROLL e um exemplo deste tipo de movimento é mostrado em 111.

Tal como mostrado na figura 1, o segundo movimento consiste na rotação da cabeça em torno do eixo y (eixo 5 responsável por modificar o olhar de cima e para baixo). Este movimento, ilustrado em 121, é conhecido como PITCH.

O terceiro movimento consiste na rotação da cabeça em torno do eixo z (responsável por modificar o olhar da esquerda para direita). Este movimento, ilustrado em 131, é 10 conhecido como YAW.

Além destes movimentos da cabeça, alguns movimentos faciais, como os das sobrancelhas, boca e olhos são importantes para tornar o avatar mais realista e gerar ações em um dispositivo portátil.

15 A. Animação do humanóide virtual

O reconhecimento dos movimentos da cabeça e feições da face forma um conjunto básico de parâmetros, capaz de identificar e reproduzir o movimento realizado por uma pessoa. Para isto, faz-se necessária a captura de imagens 20 de vídeo, o reconhecimento do movimento e reprodução em um humanóide virtual.

Uma maneira de implementar o humanóide virtual é utilizar a codificação MPEG4. Nesta implementação, existem parâmetros especializados que criam a capacidade de gerar e

transmitir o vídeo de uma "cabeça falante" sintética para serem utilizados em comunicação multimídia.

O MPEG4 inclui um conjunto de Parâmetros de Animação Facial (FAP - *Facial Animation Parameters*). Estes foram concebidos a partir do estudo de pequenas ações faciais, estando relacionados ao movimento realizado pelos músculos da face.

Atualmente, os FAPs são compostos por um conjunto de sessenta e oito parâmetros que definem a modificação da forma ou os movimentos faciais. Os FAPs possibilitam reproduzir emoções ou expressar a feição de cada sílaba de uma seqüência de áudio.

Para transmitir um vídeo adquirido a 30 quadros por segundo (fps) usando 68 parâmetros de 8 bits são necessários apenas 15,94 kbps. Com base nesta taxa, observa-se que um dispositivo celular, mesmo operando na velocidade mais baixa de transferência de dados usando o GPRS (*General Packet Radio Service*), possui a capacidade de transmitir e receber o vídeo utilizando a animação facial baseada em FAP, pois, na prática, a banda do GPRS está situada em torno dos 40 kbps, permitindo que uma comunicação *full-duplex* seja realizada.

Já nas tecnologias EDGE (*Enhanced Data rates for GSM Evolution*) e WCDMA (*Wide-Band Code-Division Multiple*

Access), a largura de banda é bem maior que a necessária para comunicação através de FAP. Nas tecnologias EDGE, a preocupação não é o funcionamento, mas sim a quantidade de dados trafegados. Os aparelhos de telefonia celular podem 5 comprimir o vídeo adquirido por sua câmera utilizando as codificações MPEG4 e H.263. Uma hora de um vídeo codificado em MPEG4 com baixa qualidade ocupa, em média, 65 Mega Bytes (MB). Por outro lado, uma hora de vídeo de FAP adquiridos a uma taxa de 15 fps gera um tráfego de apenas 28 MB. Desta 10 forma, é possível observar que há uma grande economia para o usuário final ao empregar os FAP.

B. Utilização da face como forma de interação

Uma forma mais direta de interação é realizar o movimento sem ter que pressionar qualquer componente (tecla 15 ou tela), como, por exemplo, apenas girando o dispositivo portátil. A figura 2 ilustra esta maneira de manipulação, similar a um manche. A figura 2 mostra um aparelho celular (200) posicionado na direção vertical, indicado pelo eixo 210. Ao modificar o eixo do celular para outro eixo, 20 indicado por 211 a 216, é possível gerar diferentes ações no mesmo.

Por exemplo, à medida que o ângulo do celular se afasta de 90° (eixo 210), há um acréscimo da velocidade angular na aplicação do dispositivo portátil. O sentido do

movimento é determinado pelo sentido de rotação, horário (eixos 211 a 213) e anti-horário (eixos 214 a 216). De acordo com os ensinamentos do estado da técnica, uma forma de se fazer isto é utilizar um acelerômetro integrado ao dispositivo portátil. Entretanto, esta técnica possui uma dificuldade relacionada aos custos de fabricação. Apesar de eficientes, do ponto de vista da melhoria na usabilidade do dispositivo portátil, muitos fabricantes optam por não disponibilizar este tipo de solução ao mercado devido ao aumento do valor de venda do dispositivo, causado pela inclusão desses sensores. Além disso, esta solução está restrita a novos dispositivos, pois se torna necessário redesenhar o circuito do mesmo. Uma forma de obter um efeito similar seria realizar o procedimento utilizando a câmera integrada, normalmente usada para vídeo-conferência ou como máquina fotográfica. Nesta direção, os parâmetros de pose da cabeça e algumas feições poderiam ser usados para controlar aplicações no dispositivo portátil.

Tal como ilustrado na figura 3, o dispositivo computacional pode ser composto por unidade de processamento central ou outro elemento de processamento para executar as instruções computacionais com a câmera integrada, meio para armazenamento de informações e instruções, meio de exibição, dispositivo de

entrada de informação 315, componentes de entrada e saída de áudio 316, tais como microfone e alto-falante, e componentes de conectividade a outros dispositivos computacionais através de rede com ou sem fio 317.

5 O método da presente invenção permite desde o reconhecimento da posição da cabeça do usuário, sua orientação e a variação de distância do usuário à câmera até o reconhecimento de feições como, por exemplo, o movimento das sobrancelhas e a abertura da boca.

10 O método proposto pode ser utilizado para fornecer os parâmetros para realizar a animação facial através de FAP. Além disto, através do reconhecimento da posição facial, o referido método faz com que o sistema do dispositivo móvel no qual o método está implementado possa ser utilizado como
15 dispositivo de entrada do próprio aparelho portátil e a partir disso, controlar aplicações, como por exemplo: jogos, manipulação de mapas e imagens, dentre outros. O método de rastreamento facial e reconhecimento de feições aqui descrito otimiza o esforço computacional e robustez,
20 viabilizando sua utilização em dispositivos com restrições de capacidade computacional, como os dispositivos portáteis. Este método está dividido conforme apresentado na figura 4 e compreende seguintes etapas:

1. Pré-processamento E-410, utilizando filtros de

realce espacial para suavizar as bordas da face, bem como melhorar a uniformidade das regiões segmentadas. A referida etapa E-410 é necessária devido a não uniformidade dos tons de cores da pele gerada por variações suaves de cor, existência de barba ou sinais que prejudicam a etapa de segmentação de imagens a seguir explicada.

2. Segmentação da imagem E-411 através de cor a partir da separação das regiões que possuem cores similares à cor da pele. A referida etapa E-411 utiliza um algoritmo que é baseado no método de "*Skin Color Detection*" e que nesta invenção é aplicado a imagens codificadas no formato YCbCr. Este tipo de método possui um baixo esforço computacional e consegue segmentar imagens destacando as regiões de pele de maneira eficaz para indivíduos de diferentes etnias.

3. Detecção do ângulo da cabeça E-412, a partir da imagem segmentada, calculando os momentos estatísticos da face para calcular a angulação da cabeça ou ângulo de ROLL (conforme referência 110 da figura 1), que está relacionada ao ângulo formado pelo eixo medial do rosto.

4. Remoção da rotação da cabeça E-413, através da projeção dos *pixels* da face no eixo medial do rosto para que a detecção de feições seja invariante à rotação. Desta forma, o rosto aparentemente fica sempre na mesma posição

para as etapas seguintes do processamento.

5 5. Reconhecimento dos pontos extremos da face segmentada E-414 que pertencem aos eixos centrais do rosto, com base no eixo medial do rosto, após a remoção da rotação. Estes pontos são importantes para definir o tamanho do rosto e localizar a posição dos olhos, boca, nariz e sobrancelha.

10 6. Realce de regiões faciais E-415, através de um processamento que considera a pigmentação de cor e algumas características da face, como olhos, sobrancelhas, dentre outras, que podem ser realçadas através de uma relação que envolve os canais Cb e Cr da imagem adquirida. Para fins de diminuição do esforço computacional, apenas uma região da imagem é realçada. Esta região é delimitada a partir dos
15 pontos extremos da face, evitando o cálculo deste realce fora da região de face.

20 7. Rotulação das regiões realçadas E-416, a partir da aplicação de um método de extração de elementos conectados na imagem realçada, resultando em um conjunto de regiões que separam cada tipo de objeto. Estes objetos consistem em olhos, sobrancelhas, narinas ou até mesmo pequenos ruídos causados por sinais ou manchas na pele. Em uma etapa posterior, os referidos pontos são identificados dependendo de características específicas particulares de

cada uma das estruturas (olhos, sobrancelhas e narinas).

8. Detecção de ângulos em outros eixos de rotação da cabeça (YAW e PITCH, conforme figura 1) E-417, a partir da estruturação e separação das regiões escuras próximas ao nariz feitas na etapa anterior E-416, onde os pontos pertencem a uma região próxima ao centro de massa da cabeça. Isto possibilita a determinação dos outros ângulos de rotação (YAW e PITCH) da cabeça, representados por 120 e 130 na figura 1.

9. Classificação das regiões realçadas (sobrancelhas, olhos e boca) E-418, a partir da determinação dos ângulos de rotação e utilização dos pontos extremos da face em conjunto para estimar a localização em que as referidas regiões estão dispostas na face. Com base nestas regiões, os elementos conectados separados pela etapa E-416 são classificados como sobrancelhas, olhos e boca.

10. Reconhecimento de gestos faciais E-419, a partir da extração de informações específicas de cada uma das regiões localizadas, identificando ações, como por exemplo: boca aberta ou fechada, olhos abertos ou fechados, sobrancelha normal ou levantada.

O funcionamento de um sistema com base no método de detecção e rastreamento da face e gestos faciais da

presente invenção será pormenorizado na figura 5.

Conforme mostrado na figura 5, na etapa E-510, a partir de uma seqüência de imagens obtidas pela câmera integrada ao dispositivo móvel 500, é realizada a etapa de pré-processamento E-510 através de um realce espacial feito por um filtro de média. De acordo com uma modalidade preferida da invenção, um filtro de média de tamanho 5x5 pixels gera um desempenho satisfatório e eficiente para as próximas etapas de rastreamento. Os resultados da etapa E-510 são mostrados em 511.

A etapa de segmentação da face, representada pelo bloco E-511 utiliza a informação de cor através do modelo de cor YCbCr para localizar os objetos com tonalidade de cor pré-definida.

De acordo com a modalidade preferida da invenção, o método de segmentação utilizado é a limiarização (*thresholding*) por possuir boa eficiência. Esta limiarização é definida como:

$$s(i, j) = \begin{cases} 1, & \text{se } Cb(i, j) \in [77, 127] \text{ e } Cr(i, j) \in [133, 173] \\ 0, & \text{caso contrário} \end{cases}, \quad (1)$$

em que (i, j) é a posição do pixel, $Cb(i, j)$ e $Cr(i, j)$ são as componentes de pigmentação azul e vermelha do formato de cores YCbCr, respectivamente. Desta forma, todos os pixels que possuírem os níveis de Cb e Cr dentro do intervalo

[77,127] e [133, 173], respectivamente, são considerados como pele, ou seja, $(S(i, j) = 1)$ e aqueles não pertencem ao intervalo $(S(i, j) = 0)$ são considerados como fundo.

Esquemáticamente, o resultado deste processo é
5 mostrado em 521, possuindo cor branca para o fundo e preto para as regiões segmentadas.

Após a segmentação da imagem, em E-512, é realizada a determinação do ângulo da cabeça através de momentos estatísticos da forma segmentada, tendo seu resultado
10 mostrado em 531.

O ângulo do eixo medial (ângulo de ROLL) da cabeça é determinado a partir dos momentos estatísticos. Nesta formulação matemática é necessário fazer o cálculo de uma função arco-tangente. Entretanto, esta função requer um
15 custo computacional elevado. Assim sendo, de acordo com a concretização preferida da invenção, a referida função foi implementada utilizando uma tabela de pesquisa de 180 posições. O ângulo da cabeça corresponde ao valor de tangente mais próxima à encontrada nesta tabela.

20 Após a localização do contorno a ser rastreado, a etapa E-513 desfaz a rotação, que tem seu resultado mostrado em 541.

Após isto, na etapa E-514, é realizada a localização dos pontos extremos da face, que são pontos que pertencem

ao eixo medial da face.

A figura 6 ilustra a etapa de localização dos pontos extremos. Na figura 6, a referência 600 apresenta uma face que passou pela correção de inclinação. Isto é, a face tem sua orientação bem próxima da vertical, sem rotação. As coordenadas (X', Y') tem como origem o centro de massa da face, representado pela referência 610. Idealmente, todos os pontos pertencentes ao eixo medial horizontal 611 possuem coordenada Y' igual a zero e os pontos pertencentes ao eixo medial vertical 612 possuem X' igual a zero. Desta forma, os pontos extremos 621, 622, 623 e 624, nomeados como P_1 , P_2 , P_3 e P_4 possuem as coordenadas $(X_{P_1}, 0)$, $(X_{P_2}, 0)$, $(0, Y_{P_3})$ e $(0, Y_{P_4})$. Considerando a aproximação feita para cálculo do ângulo da cabeça, conforme acima mencionado, como ilustrado na figura 6, estabelece-se uma faixa de tolerância representada por 630 e 640. Isto faz com que o ponto extremo seja localizado mesmo com o ângulo da cabeça aproximado, conforme ilustrado por 631, 632, 641 e 642.

Em algumas situações, uma parte do pescoço é também detectada durante o processo de segmentação da pele. Para evitar esse inconveniente, no cálculo do ponto P_3 é utilizada uma restrição, na qual a distância entre os extremos horizontais da face é considerada. Assim, o ponto

P3 não pode estar localizado a uma distância maior do que o ponto médio vertical e P4; e a distância horizontal P2 e P3 multiplicada por um fator de escala N. De acordo com a modalidade aqui preferida, o referido fator N deve ser um
 5 número real contido no intervalo de 1,5 a 2, preferencialmente 1,7.

Após a determinação dos pontos extremos, é realizado um processamento de realce de regiões que representam a pupila, as narinas e as sobrancelhas E-515. Através de uma
 10 combinação dos canais do formato de cor YCbCr, é obtida a seguinte expressão para realce:

$$R(i, j) = \frac{255 \cdot Cb(i, j)}{Cr(i, j) + 1} - 2 \cdot Y(i, j), \quad (2)$$

em que (i, j) é a posição do pixel e $Y(i, j)$, $Cb(i, j)$ e $Cr(i, j)$ são os componentes de luminância, pigmentação azul
 15 e vermelha do formato de cores YCbCr, respectivamente. Após o processamento, os valores realçados são normalizados entre 0 e 1 e se aplica uma limiarização na imagem realçada. De maneira experimental, foi determinado pelos inventores da presente invenção que os pontos que
 20 satisfazem a expressão lógica:

$$\left[(S(i, j) = 1) \& \left(\frac{R(i, j)}{\max_{v_i, j}(R(i, j))} > 0,706 \right) \right], \quad (3)$$

são considerados como olhos, sobrancelhas e boca, enquanto que os que obedecem a inequação:

$$\frac{R(i,j)}{\max_{\forall i,j}(R(i,j))} > 0,667, \quad (4)$$

são considerados como narinas.

5 O processamento não é feito em toda imagem, mas apenas em uma pequena região formada pelos quatro pontos extremos localizados pela etapa anterior, conforme mostrado na figura 7. Na figura 7, é mostrado um retângulo 700, cujos cantos extremos superior e inferior 710 e 720 possuem
10 coordenadas (X_{P1}, Y_{P3}) e (X_{P2}, Y_{P4}) , respectivamente.

A limiarização apenas separa regiões de interesse em uma imagem. Para classificar os gestos faciais, é necessário distinguir os objetos baseados na conectividade de *pixels*. Para isto, utiliza-se um algoritmo de extração
15 de elementos conectados, também conhecido como algoritmo de rotulamento. Este tipo de algoritmo consiste em percorrer a imagem *pixel-a-pixel*, dando etiquetas iguais a *pixels* conectados entre si.

Na figura 5, o resultado de etapa E-516 é mostrado em
20 571. Como ilustrado na figura 8, cada uma das regiões da face é rotulada com um número. No exemplo mostrado na figura 5, referência 571, o olho direito é rotulado com o número 1, o olho esquerdo como 2, o nariz como 3 e a boca

como 4.

Após a rotulação, na etapa E-517, é realizada a determinação do ângulo de YAW, conforme mostrado em 581. As narinas são as regiões ressaltadas mais próximas ao centro de massa da cabeça. Para localizar a posição das narinas (X_N, Y_N) utiliza-se a seguinte expressão:

$$(x_N, y_N) | N = \min_{n \in (R=k)} (|x_k - x_{CM}| + |y_k - y_{CM}| | |y_i \leq \delta|), \quad (5)$$

em que x_k é o centro de massa de uma região numerada como k durante a etapa de rotulação das regiões realçadas.

As narinas se localizam na região 641, como mostrado na figura 6, durante o movimento de rotação da esquerda para direita (YAW). Encontrada a região das narinas, a seguinte equação é usada para calcular o ângulo ω de YAW

$$\omega = \frac{200 \cdot (x_N - x_{CM})}{|x_{P1} - x_{P2}|} \quad (6)$$

O valor de ω varia entre -100 e 100, correspondendo às inclinações máximas para à direita e para à esquerda, respectivamente.

Similarmente, a determinação do ângulo ϕ de PITCH pode ser feita através de:

$$\phi = \frac{200 \cdot (y_N - y_{CM})}{|y_{P4} - y_{P3}|} \quad (7)$$

Em seguida, as sobrancelhas, olhos e boca são

localizados na etapa E-518. A figura 8 ilustra a posição destas regiões em função dos pontos extremos P1 a P4, representados por 621 a 624. A região 810 localiza o olho direito do usuário. Ela é uma região quadrada delimitada pelos cantos 811 e 812 cujas respectivas coordenadas (X_{OD}, Y_{OD}) e (X'_{OD}, Y'_{OD}) são calculadas por:

$$X_{OD} = X_{P1} + 0,4 \cdot (X_{P2} - X_{P1}) + d \sin \omega , \quad (8)$$

$$Y_{OD} = Y_{P4} - 0,25 \cdot (X_{P4} - X_{P3}) , \quad (9)$$

$$X'_{OD} = X_{P1} + 0,2 \cdot (X_{P2} - X_{P1}) + d \sin \omega , \quad (10)$$

$$10 \quad Y'_{OD} = Y_{P4} - 0,45 \cdot (X_{P4} - X_{P3}) . \quad (11)$$

Da mesma forma, a região 820 localiza o olho esquerdo do usuário. Ela é uma região quadrada delimitada pelos cantos 821 e 822 cujas respectivas coordenadas (X_{OE}, Y_{OE}) e (X'_{OE}, Y'_{OE}) são calculadas por:

$$15 \quad X_{OE} = X_{P2} - 0,4 \cdot (X_{P2} - X_{P1}) + d \sin \omega , \quad (12)$$

$$Y_{OE} = Y_{P4} - 0,25 \cdot (X_{P4} - X_{P3}) , \quad (13)$$

$$X'_{OE} = X_{P2} - 0,2 \cdot (X_{P2} - X_{P1}) + d \sin \omega , \quad (14)$$

$$Y'_{OE} = Y_{P4} - 0,45 \cdot (X_{P4} - X_{P3}) , \quad (15)$$

e a região 830 localiza a boca do usuário. Ela é uma região quadrada delimitada pelos cantos 831 e 832 cujas respectivas coordenadas (X_B, Y_B) e (X'_B, Y'_B) são calculadas por:

$$X_B = X_{P3} - 0,25 \cdot (X_{P2} - X_{P1}) + d \sin \omega , \quad (16)$$

$$Y_B = Y_{P4} - 0,62 \cdot (Y_{P4} - Y_{P3}) , \quad (17)$$

$$X'_B = X_{P3} + 0,25 \cdot (X_{P2} - X_{P1}) + d \sin \omega , \quad (18)$$

$$Y'_B = Y_{P4} - 0,92 \cdot (Y_{P4} - Y_{P3}) . \quad (19)$$

5 O parâmetro d é um fator para correção da posição, ocorrendo uma variação do ângulo de YAW, fazendo com que o método ajuste a posição da janela de acordo com o ângulo de rotação da cabeça.

Após a conclusão das etapas E-510 a E-518, é possível
10 analisar o comportamento de cada uma das referidas regiões e detectar os gestos realizados E-519.

De acordo com a presente invenção, para os olhos e para a boca são considerados dois gestos: boca fechada ou aberta. Para isto é realizada uma comparação entre a área
15 da boca e a área da região em que se localiza a boca. Quando esta é pequena, significa que a boca está fechada e quando a área é maior do que um certo limiar, a boca é considerada como aberta. Da mesma forma é feito para os olhos, porém com a sobrancelha o processo é diferente.

20 A figura 9 mostra pontos importantes da sobrancelha. Em 900, é mostrada a sobrancelha. Para definir o gesto feito com ela, detectam-se os pontos extremos da

sobrancelha mais à esquerda S1, mais à direita S2 e o extremo mais acima e próximo ao centro de massa S3. A figura 9 mostra os três gestos possíveis de reconhecimento, de acordo com os ensinamentos da presente invenção. O primeiro gesto 900 representa a sobrancelha normal. O

5 segundo gesto 910, a sobrancelha levantada ao centro e o terceiro 920, a sobrancelha com a extremidade levantada. A detecção destes gestos é realizada através das distâncias horizontais entre os pontos calculados por: $D_1 = |x_{S_1} - x_{S_2}|$,

10 $D_2 = |x_{S_1} - x_{S_3}|$ e $D_3 = |x_{S_2} - x_{S_3}|$.

Através das seguintes regras, os gestos feitos com as sobrancelhas são reconhecidos

1. Sobrancelha normal: quando as distâncias D_1 , D_2 e D_3 são menores que um limiar de tolerância Δ , ou seja,

15
$$D_1 < \Delta \text{ e } D_2 < \Delta \text{ e } D_3 < \Delta. \quad (20)$$

2. Sobrancelha centro levantada: quando a distância D_1 é menor que o limiar de tolerância, enquanto D_2 e D_3 são maiores que o mesmo limiar, ou seja,

$$D_1 < \Delta \text{ e } D_2 > \Delta \text{ e } D_3 > \Delta. \quad (21)$$

20 3. Sobrancelha extremo levantada: quando a distância D_3 é menor que o limiar de tolerância, enquanto D_1 e D_2 são maiores que o mesmo limiar, ou seja,

$$D_1 > \Delta \text{ e } D_2 > \Delta \text{ e } D_3 < \Delta. \quad (22)$$

O método aqui descrito ainda pode ser implementado, de forma genérica, em dispositivos computacionais, como por exemplo: um computador pessoal que possua componente de
5 captura de imagem, tal como uma *webcam* ou dispositivos portáteis como PDAs, celulares e *Smart Phones* que possuam câmara digital.

A presente invenção tendo sido descrita vai ser evidente para uma pessoa versada na técnica que muitas
10 alterações e mudanças podem ser feitas na mesma, sem que se afaste do espírito ou do escopo da referida invenção, como definido nas reivindicações anexas.

REIVINDICAÇÕES

1) Método de rastreamento facial e identificação de feições em dispositivos portáteis **caracterizado pelo** fato de que as etapas de:

5 ➤ pré-processamento (E-410), utilizando filtros de realce espacial para suavizar as bordas da face e melhorar a uniformidade das regiões segmentadas;

 ➤ segmentação da imagem (E-411) através de cor a partir da separação das regiões que possuem cores similares
10 à cor da pele;

 ➤ detecção do ângulo da cabeça (E-412), a partir da imagem segmentada, calculando os momentos estatísticos da face para calcular a angulação da cabeça ou ângulo de ROLL (110), que está relacionada ao ângulo formado pelo eixo
15 medial do rosto;

 ➤ remoção da rotação da cabeça (E-413), através da projeção dos *pixels* da face no eixo medial do rosto para que a detecção de feições seja invariante à rotação;

 ➤ reconhecimento dos pontos extremos da face
20 segmentada (E-414) que pertencem aos eixos centrais do rosto, com base no eixo medial do rosto, após a remoção da rotação;

 ➤ realce de regiões faciais (E-415), através de um

processamento que considera a pigmentação de cor e algumas características da face, como olhos, sobrancelhas, dentre outras, que podem ser realçadas através de uma relação que envolve os canais Cb e Cr da imagem adquirida;

5 ➤ rotulação das regiões realçadas (E-416), a partir da aplicação de um método de extração de elementos conectados na imagem realçada, resultando em um conjunto de regiões que separam cada tipo de objeto;

10 ➤ detecção de ângulos em outros eixos de rotação da cabeça (E-417), a partir da estruturação e separação das regiões escuras próximas ao nariz feitas na etapa anterior (E-416), onde os pontos pertencem a uma região próxima ao centro de massa da cabeça, possibilitando a determinação dos outros ângulos de rotação da cabeça (120) e (130);

15 ➤ classificação das regiões realçadas (E-418), a partir da determinação dos ângulos de rotação e utilização dos pontos extremos da face em conjunto para estimar a localização em que as referidas regiões estão dispostas na face;

20 ➤ reconhecimento de gestos faciais (E-419), a partir da extração de informações específicas de cada uma das regiões localizadas, identificando as ações da face.

2) Método, de acordo com a reivindicação 1, **caracterizado pelo** fato de que as referidas ações

identificadas na etapa de reconhecimento de gestos faciais (E-419) são boca aberta ou fechada, olhos abertos ou fechados, sobrancelha normal ou levantada, dentre outras.

3) Método, de acordo com a reivindicação 1,
5 **caracterizado pelo** fato de que a referida etapa de segmentação da imagem (E-411) segmenta as imagens destacando as regiões de pele de maneira eficaz para indivíduos de diferentes etnias.

4) Método, de acordo com a reivindicação 1,
10 **caracterizado pelo** fato de que os referidos pontos extremos identificados na referida etapa de reconhecimento dos pontos extremos da face segmentada (E-414) definem o tamanho do rosto e localizam a posição dos olhos, boca, nariz e sobrancelha.

15 5) Método, de acordo com a reivindicação 1, **caracterizado pelo** fato de que, para a diminuição do esforço computacional, na referida etapa de realce de regiões faciais (E-415), apenas uma região da imagem é realçada, a referida região sendo delimitada a partir dos
20 referidos pontos extremos da face, evitando o cálculo do referido realce fora da região da face.

6) Método, de acordo com a reivindicação 1, **caracterizado pelo** fato de que a referida etapa de pré-processamento E-510 é realizada através de um realce

espacial feito por um filtro de média, preferencialmente, um filtro de média de tamanho 5x5 pixels.

7) Método, de acordo com a reivindicação 1, **caracterizado pelo** fato de que a referida etapa de segmentação da imagem (E-411) utiliza a informação de cor através do modelo de cor YCbCr para localizar os objetos com tonalidade de cor pré-definida.

8) Método, de acordo com a reivindicação 1, **caracterizado pelo** fato de que a referida etapa de segmentação da imagem (E-411) utiliza a limiarização que é definida como:

$$s(i, j) = \begin{cases} 1, & \text{se } Cb(i, j) \in [77, 127] \text{ e } Cr(i, j) \in [133, 173] \\ 0, & \text{caso contrário} \end{cases}$$

onde (i, j) é a posição do pixel, $Cb(i, j)$ e $Cr(i, j)$ são as componentes de pigmentação azul e vermelha do formato de cores YCbCr, respectivamente.

9) Método, de acordo com a reivindicação 8, **caracterizado pelo** fato de que todos os pixels que possuírem os níveis de Cb e Cr dentro do intervalo $[77, 127]$ e $[133, 173]$, respectivamente, são considerados como pele, ou seja, $(S(i, j) = 1)$ e aqueles não pertencem ao intervalo $(S(i, j) = 0)$ são considerados como fundo.

10) Método, de acordo com a reivindicação 1, **caracterizado pelo** fato de que na referida etapa de

reconhecimento dos pontos extremos da face segmentada (E-414) estabelece-se uma faixa de tolerância (630) e (640), possibilitando a localização do ponto extremo mesmo com o ângulo da cabeça aproximado.

5 11) Método, de acordo com a reivindicação 10, **caracterizado pelo** fato de que na referida etapa de reconhecimento dos pontos extremos da face segmentada (E-414), para o cálculo do ponto P3, é utilizada uma restrição, na qual a distância entre os extremos
10 horizontais da face é considerada.

 12) Método, de acordo com a reivindicação 11, **caracterizado pelo** fato de que o referido ponto P3 não está localizado a uma distância maior do que o ponto médio vertical e P4; e a distância horizontal P2 e P3
15 multiplicada por um fator de escala N.

 13) Método, de acordo com a reivindicação 12, **caracterizado pelo** fato de que o referido fator N é um número real contido no intervalo de 1,5 a 2.

 14) Método, de acordo com a reivindicação 13,
20 **caracterizado pelo** fato de que o referido fator N é preferencialmente 1,7.

 15) Método, de acordo com a reivindicação 1, **caracterizado pelo** fato de que a referida etapa de realce de regiões faciais (E-415) é calculada a partir da seguinte

expressão:

$$R(i, j) = \frac{255 \cdot Cb(i, j)}{Cr(i, j) + 1} - 2 \cdot Y(i, j),$$

em que (i, j) é a posição do pixel e $Y(i, j)$, $Cb(i, j)$ e $Cr(i, j)$ são os componentes de luminância, pigmentação azul e vermelha do formato de cores YCbCr, respectivamente.

16) Método, de acordo com a reivindicação 1, **caracterizado pelo** fato de que os pontos que satisfazem a expressão lógica:

$$\left[(S(i, j) = 1) \& \& \left(\frac{R(i, j)}{\max_{\forall i, j} (R(i, j))} > 0,706 \right) \right],$$

10 são considerados como olhos, sobrancelhas e boca, enquanto que os que obedecem a inequação:

$$\frac{R(i, j)}{\max_{\forall i, j} (R(i, j))} > 0,667,$$

são considerados como narinas.

17) Método, de acordo com a reivindicação 16, **caracterizado pelo** fato de que para localizar a posição das narinas (X_N, Y_N) utiliza-se a seguinte expressão:

$$(x_N, y_N) | N = \min_{n \in (R=k)} (|x_k - x_{CM}| + |y_k - y_{CM}| | |y_i \leq \delta|),$$

em que x_k é o centro de massa de uma região numerada como k durante a etapa de rotulação das regiões realçadas.

18) Método, de acordo com a reivindicação 1,

caracterizado pelo fato de que após a localização da região das narinas, a seguinte equação é usada para calcular o ângulo ω de YAW:

$$\omega = \frac{200 \cdot (x_N - x_{CM})}{|x_{P1} - x_{P2}|},$$

5 o valor de ω variando entre -100 e 100, correspondendo às inclinações máximas para à direita e para à esquerda, respectivamente.

19) Método, de acordo com a reivindicação 1, **caracterizado pelo** fato de que a determinação do ângulo φ de PITCH é feita através da seguinte expressão:

$$\varphi = \frac{200 \cdot (y_N - y_{CM})}{|y_{P1} - y_{P2}|}.$$

20) Método, de acordo com a reivindicação 1, **caracterizado pelo** fato de que o olho direito do usuário é uma região quadrada delimitada pelos cantos (811) e (812) cujas respectivas coordenadas (X_{OD}, Y_{OD}) e (X'_{OD}, Y'_{OD}) são calculadas por:

$$X_{OD} = X_{P1} + 0,4 \cdot (X_{P2} - X_{P1}) + d \sin \omega,$$

$$Y_{OD} = Y_{P4} - 0,25 \cdot (X_{P4} - X_{P3}),$$

$$X'_{OD} = X_{P1} + 0,2 \cdot (X_{P2} - X_{P1}) + d \sin \omega,$$

$$Y'_{OD} = Y_{P4} - 0,45 \cdot (X_{P4} - X_{P3}).$$

20

21) Método, de acordo com a reivindicação 1, **caracterizado pelo** fato de que as coordenadas para o olho esquerdo são calculadas a partir a partir das seguintes expressões:

$$5 \quad X_{OE} = X_{P2} - 0,4 \cdot (X_{P2} - X_{P1}) + d \sin \omega ,$$

$$Y_{OE} = Y_{P4} - 0,25 \cdot (X_{P4} - X_{P3}) ,$$

$$X'_{OE} = X_{P2} - 0,2 \cdot (X_{P2} - X_{P1}) + d \sin \omega ,$$

$$Y'_{OE} = Y_{P4} - 0,45 \cdot (X_{P4} - X_{P3})$$

22) Método, de acordo com a reivindicação 1, **caracterizado pelo** fato de que as coordenadas para a boca são calculadas a partir das seguintes expressões:

$$X_B = X_{P3} - 0,25 \cdot (X_{P2} - X_{P1}) + d \sin \omega ,$$

$$Y_B = Y_{P4} - 0,62 \cdot (Y_{P4} - Y_{P3}) ,$$

$$X'_B = X_{P3} + 0,25 \cdot (X_{P2} - X_{P1}) + d \sin \omega ,$$

$$15 \quad Y'_B = Y_{P4} - 0,92 \cdot (Y_{P4} - Y_{P3}) .$$

23) Método, de acordo com as reivindicações 20, 21 e 22, **caracterizado pelo** fato de que o parâmetro d é um fator para correção da posição, ocorrendo uma variação do ângulo de YAW, fazendo com que o método ajuste a posição da janela de acordo com o ângulo de rotação da cabeça.

24) Método, de acordo com a reivindicação 1,

caracterizado pelo fato de que a detecção de gestos da sobancelha é realizada através das distâncias horizontais entre os pontos calculados por: $D_1 = |x_{s_1} - x_{s_2}|$, $D_2 = |x_{s_1} - x_{s_3}|$ e $D_3 = |x_{s_2} - x_{s_3}|$.

- 5 25) Método, de acordo com a reivindicação 24, **caracterizado pelo** fato de que os gestos feitos com as sobancelhas são reconhecidos como: sobancelha normal: quando as distâncias D_1 , D_2 e D_3 são menores que um limiar de tolerância Δ , ou seja, $D_1 < \Delta$ e $D_2 < \Delta$ e $D_3 < \Delta$;
- 10 sobancelha centro levantada: quando a distância D_1 é menor que o limiar de tolerância, enquanto D_2 e D_3 são maiores que o mesmo limiar, ou seja, $D_1 < \Delta$ e $D_2 > \Delta$ e $D_3 > \Delta$; sobancelha extremo levantada: quando a distância D_3 é menor que o limiar de tolerância, enquanto D_1 e D_2 são maiores que o
- 15 mesmo limiar, ou seja, $D_1 > \Delta$ e $D_2 > \Delta$ e $D_3 < \Delta$.

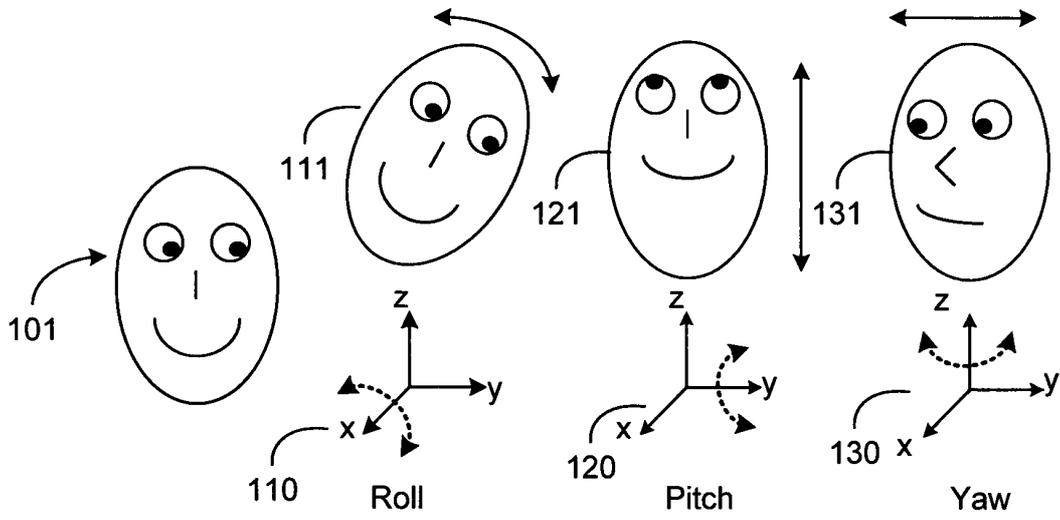


FIG 1

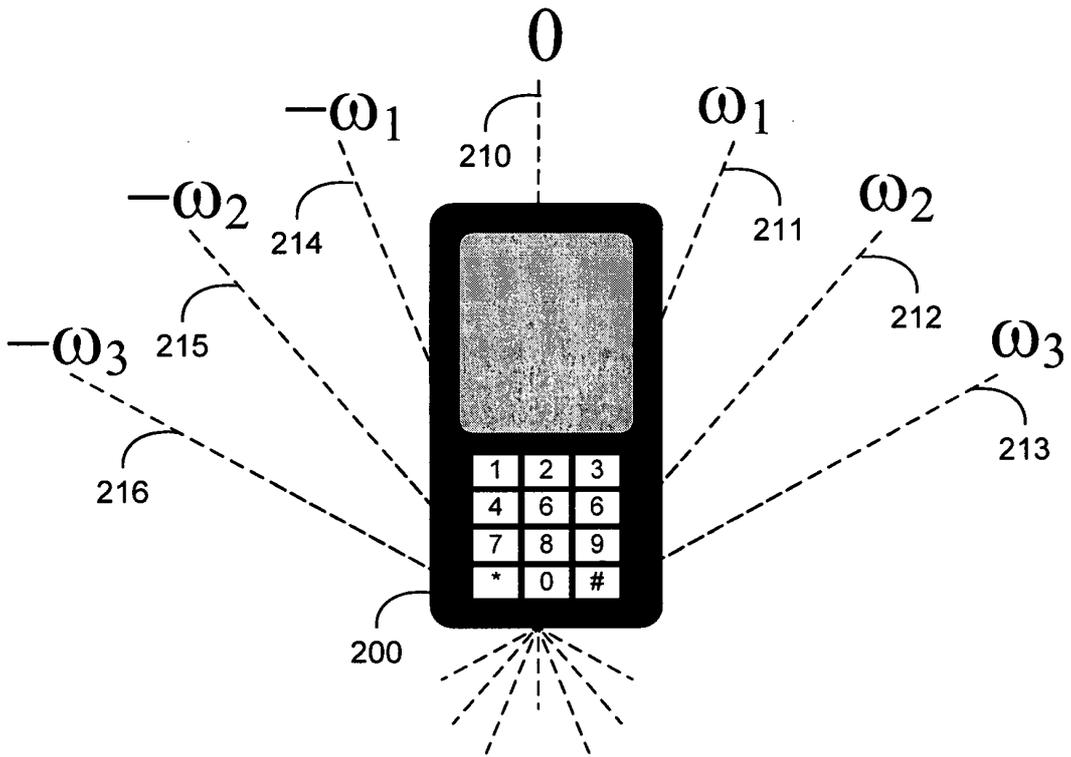


FIG 2

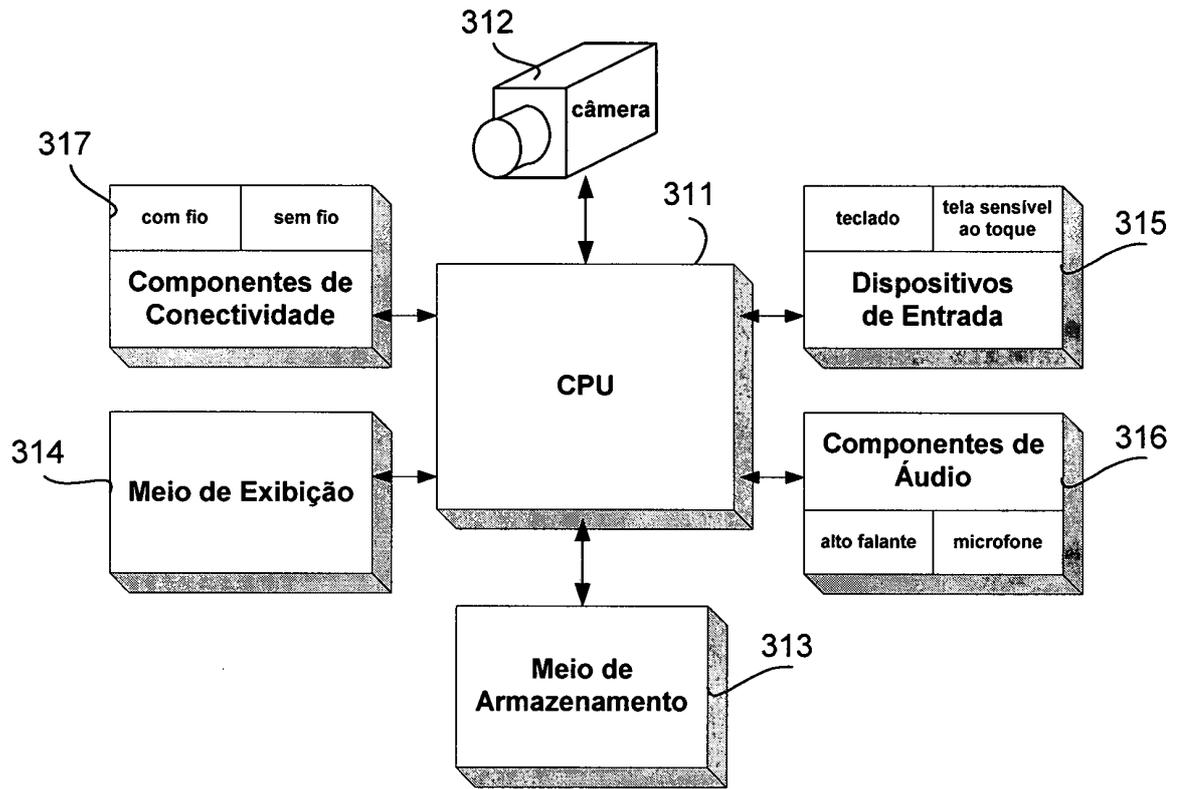
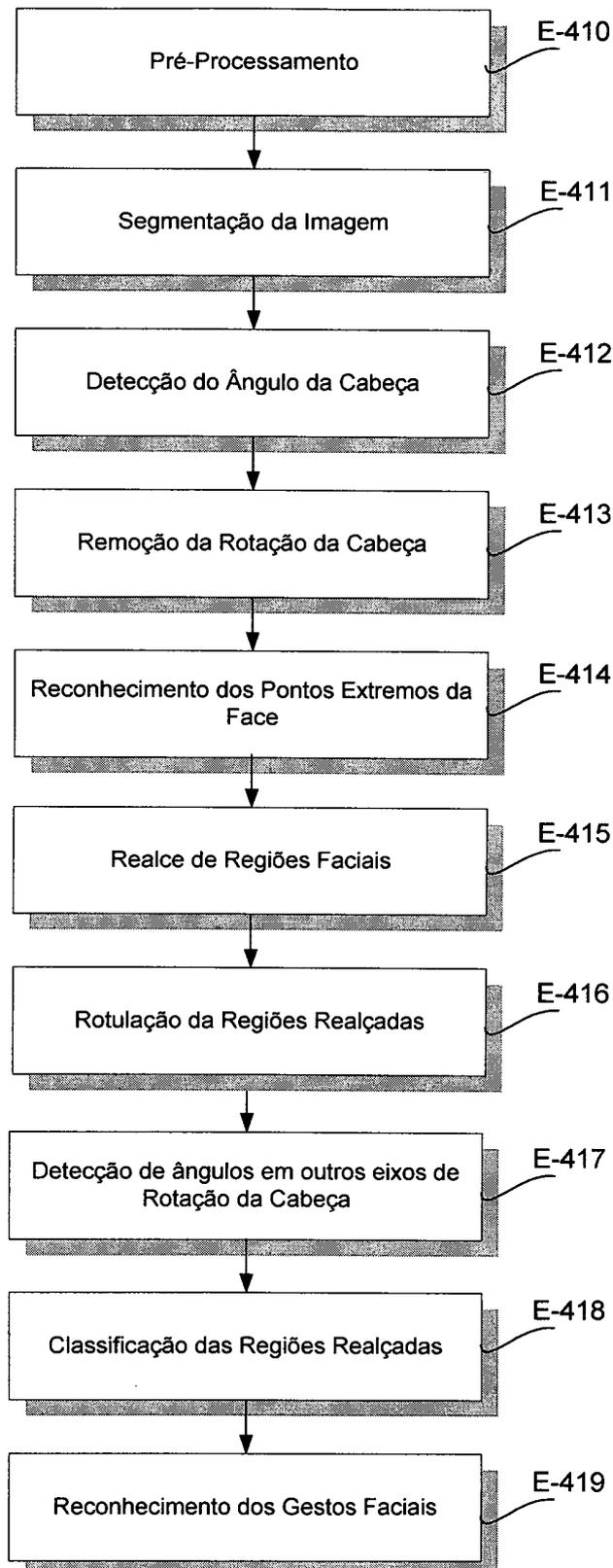


FIG 3

**FIG 4**

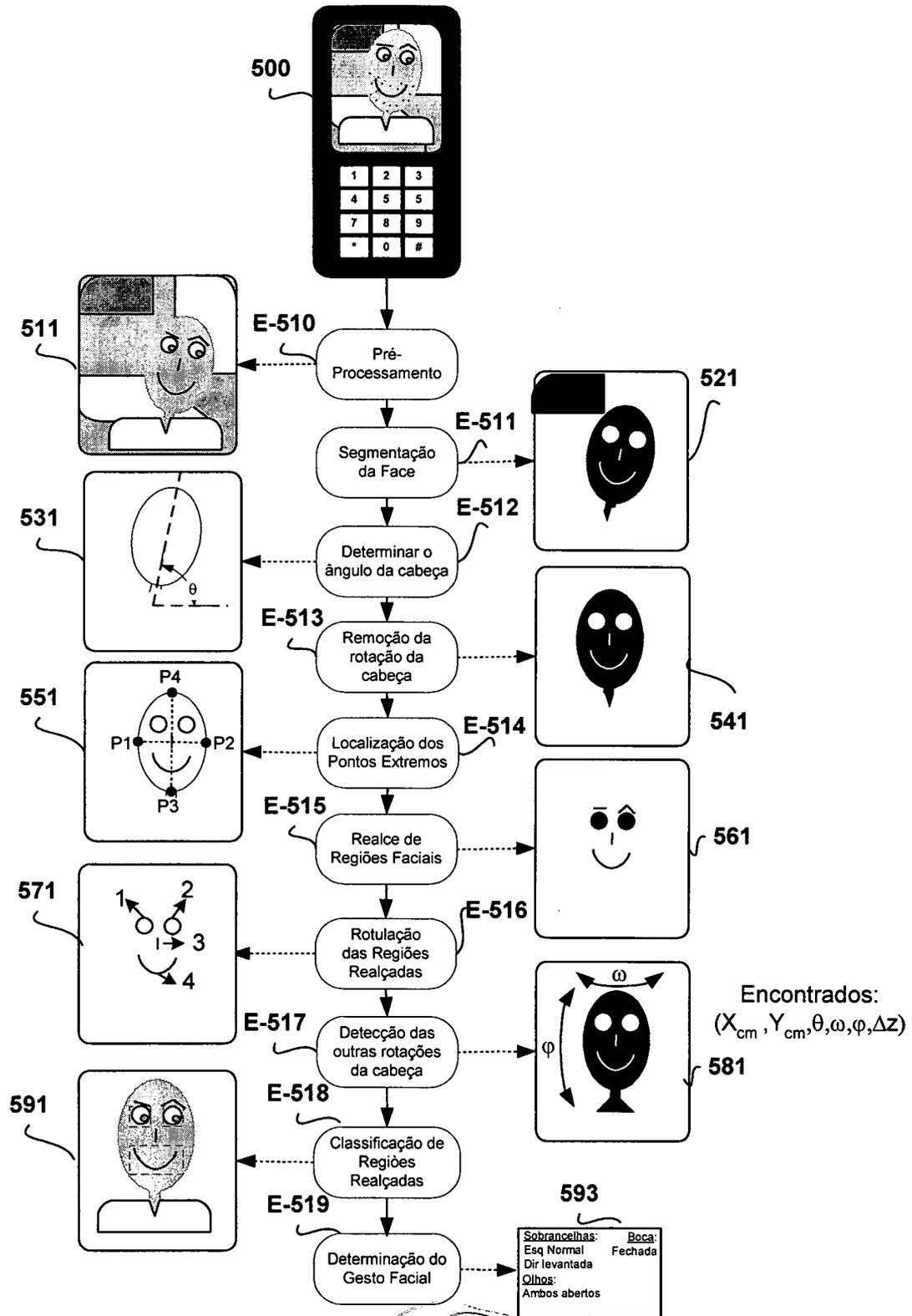


FIG 5

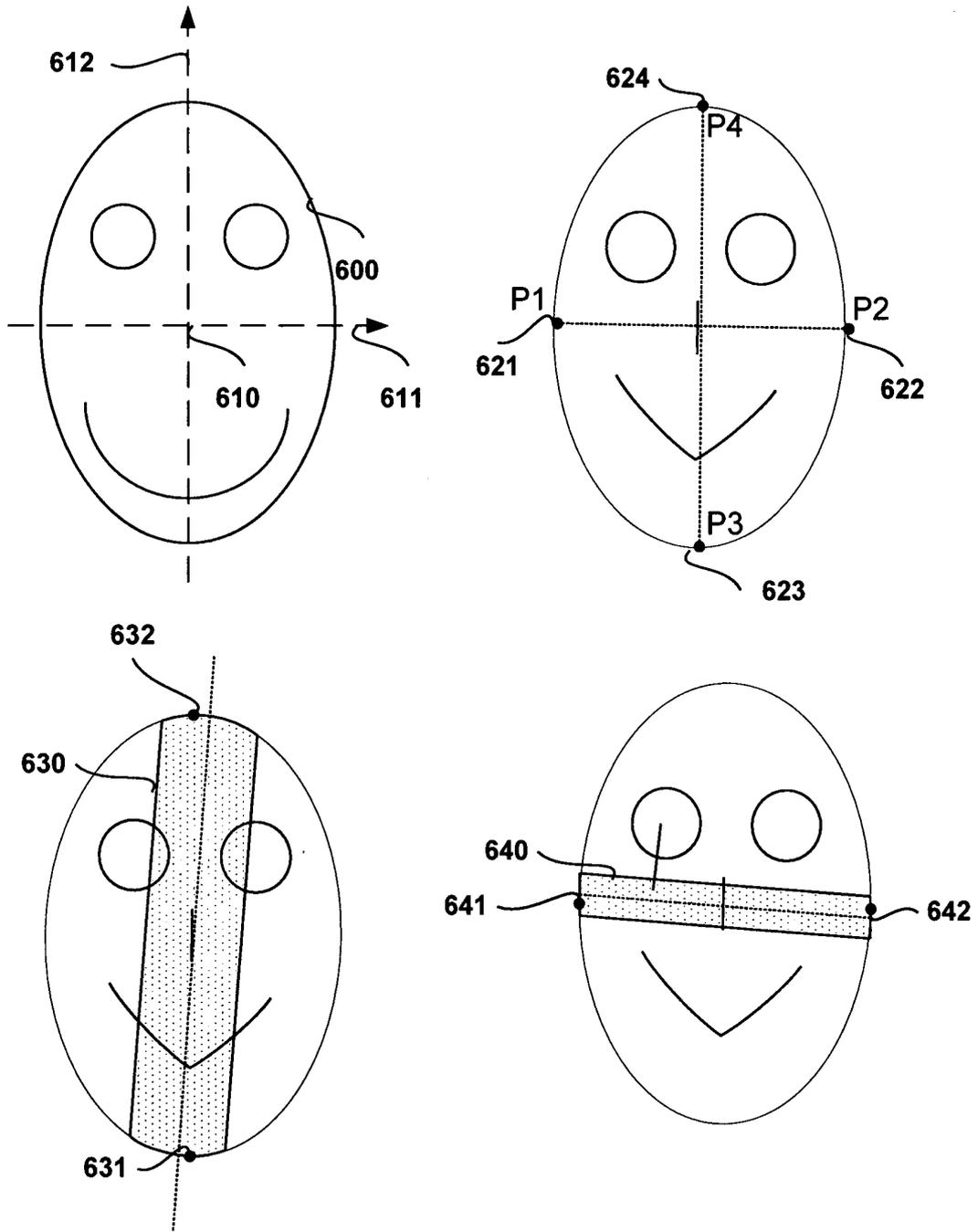


FIG 6

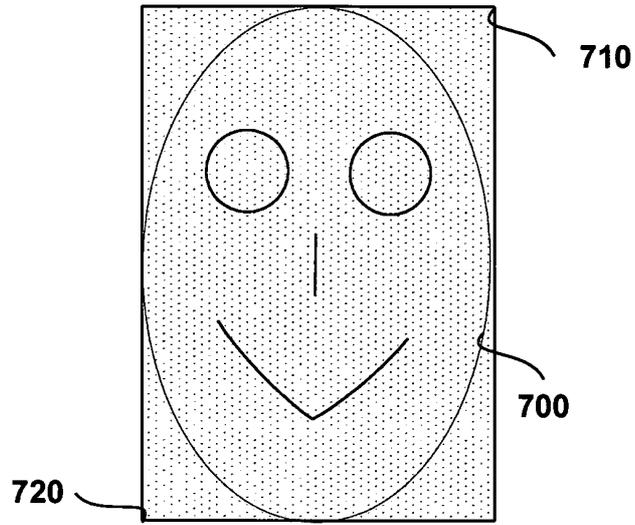


FIG 7

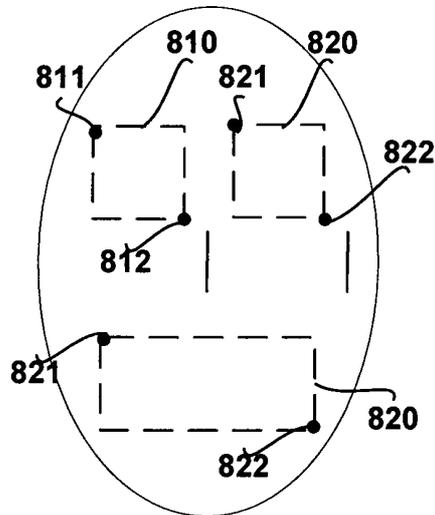


FIG 8

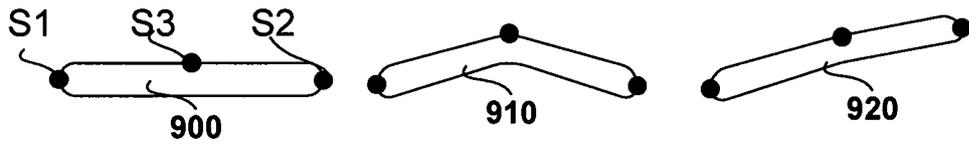


FIG 9

Resumo da Patente de Invenção para: **"MÉTODO DE RASTREAMENTO FACIAL E IDENTIFICAÇÃO DE FEIÇÕES EM DISPOSITIVOS PORTÁTEIS"**.

A presente invenção diz respeito a um método de
5 detecção e rastreamento da face e gestos faciais a partir
de técnicas de segmentação de imagens. O método da presente
invenção visa fornecer uma forma prática para interagir com
dispositivos portáteis através da utilização de movimentos
da cabeça. O método envolve a captura de uma seqüência de
10 imagens, a segmentação da face existente nas imagens, o
rastreamento dos movimentos realizados pela cabeça e
reconhecimento de expressões faciais. O método proposto é
capaz de controlar aplicações no próprio dispositivo
portátil, simulando um controle virtual, em que os comandos
15 são executados de acordo com os gestos e feições faciais
reconhecidas. Isto permite o controle da navegação de menus
e jogos, dentre outras aplicações.