

# ITALIAN PATENT OFFICE

Document No.

102011901978959A1

Publication Date

20130315

Applicant

STMICROELECTRONICS S.R.L.

Title

SISTEMA E PROCEDIMENTO PER TAGGARE CONTENUTI MULTIMEDIALI, E  
RELATIVO PRODOTTO INFORMATICO

**DESCRIZIONE** dell'invenzione industriale dal titolo:

"Sistema e procedimento per taggare contenuti multimediali,  
e relativo prodotto informatico"

di: STMicroelectronics S.r.l., di nazionalità italiana, via  
C. Olivetti 2 - 20041 Agrate Brianza (MB) - Italia

Inventori designati: Marco PESSIONE, Alexandro SENTINELLI

Depositato in data: 15 settembre 2011.

\*\*\*

### **TESTO DELLA DESCRIZIONE**

#### Campo

La descrizione riguarda la codifica di semantica multimediale.

Varie forme di realizzazione possono riguardare l'implementazione della taggatura automatica di metadati.

#### Sfondo tecnologico

Vari documenti descrivono il reperimento di informazioni multimediali, o MIR. In quest'ambito una tag viene prodotta mentre si scatta una fotografia o si gira un video. La stessa tag può essere utilizzata per reperire quel contenuto (solitamente memorizzato come file). Questo tipo di interrogazione reperimento è semplicemente denominata "interrogazione testo".

O mediante le tradizionali tag di interrogazione testo (come quelle utilizzate nei motori di ricerca, quali "Google"), o mediante interrogazioni più sofisticate basate su estrazione delle caratteristiche, il reperimento può ancora essere basato su un semplice matching dell'interrogazione e dei metadati del contenuto. I metadati sono uniti o estratti a/da un contenuto multimediale, ma ancora legati a quei particolari dati.

Un algoritmo hash protetto, che è divenuto popolare con le applicazioni di condivisione file di Torrent, può essere utilizzato per proteggere l'integrità dei dati. In uno scenario tipico, un utente U1 prima può produrre un hash H1 che è una sorta di firma digitale dell'integrità dei dati D1. L'utente U1 invia ad un altro utente U2 i dati D1 e l'hash H1. Dopo aver ricevuto i dati D2, l'utente U2 può produrre il proprio hash H2 e verificare la corrispondenza con H1; se la corrispondenza è verificata, si ritiene che D2 sia uguale a D1.

Negli ultimi vent'anni, l'evoluzione dei microcomputer ha avuto un enorme impatto sullo sviluppo della strumentazione medica. In questo settore, l'accresciuto potere di elaborazione e la capacità di tale potere di essere compattato in chip relativamente piccoli ha reso possibile la creazione di dispositivi "intelligenti" che possono essere adattati ad uno specifico paziente. Tali dispositivi possono essere in grado di monitorare, rilevare e riconoscere problemi specifici per un paziente durante la normale vita quotidiana. Monitor indossabili possono pertanto essere sviluppati grazie ai miglioramenti in termini di riduzione delle dimensioni, dei costi e del consumo energetico. Disponendo di tali monitor indossabili, che arrecano poco disturbo al paziente, è possibile memorizzare informazioni sul paziente e trasmetterle ad un ospedale locale utilizzando una rete di telecomunicazioni. Questo miglioramento va a beneficio sia del paziente, dal momento che può tornare a casa il prima possibile, mentre è ancora in fase di monitoraggio, e dell'ospedale, in quanto si può risparmiare denaro e i letti possono essere liberati per nuovi pazienti.

Diversi studi sono già stati dedicati alla stima di

come un paziente "si sente", principalmente allo scopo di trovare una correlazione tra lo stato emotivo e gli indici psicologici dell'utente.

Ora verranno discussi alcuni documenti relativi a tali studi. Questi documenti saranno indicati con un numero di riferimento tra parentesi quadre (ossia [X]), il numero riferendosi alla lista riprodotta alla fine della presente descrizione.

Kim et al. [1] hanno sviluppato un sistema di riconoscimento delle emozioni basato su segnali fisiologici, combinante l'elettrocardiogramma, la variazione della temperatura della pelle e i segnali dell'attività elettrodermica. Dopo l'elaborazione e l'estrazione delle caratteristiche, mediante l'utilizzo di una macchina/classificatore a vettori di supporto, un tale sistema permette di classificare quattro caratteristiche emozione-specifiche differenti.

Wagner et al. [2] hanno proposto un sistema di riconoscimento delle emozioni che comprende un'analisi dei dati e una classificazione dell'elettromiogramma, dell'elettrocardiogramma, della conduttività della pelle e delle variazioni della respirazione e che si avvale di un metodo di induzione musicale che suscita reazioni emotive naturali nel soggetto.

Goldstein et al. [3] hanno condotto uno studio in cui la pressione sanguigna (sia sistolica che diastolica) era correlata con diversi tipi di stati di rabbia.

Shapiro et al. [4] hanno analizzato la relazione tra l'intensità dei singoli stati d'animo e le combinazioni degli stessi misurando la pressione sanguigna e la frequenza cardiaca nelle infermiere, e hanno sperimentato un graduale aumento nella pressione sanguigna e nella

frequenza cardiaca con un maggiore indice di stati d'animo negativi e una diminuzione delle stesse per uno stato d'animo collegato col livello energetico.

#### Scopo e sintesi

Uno scopo della presente divulgazione è quindi quello di provvedere migliori forme di realizzazione rispetto alle disposizioni di cui sopra.

Secondo l'invenzione, tale scopo è raggiunto mediante un sistema che presenta le caratteristiche descritte più avanti nelle rivendicazioni che seguono.

L'invenzione riguarda inoltre un corrispondente procedimento implementato mediante elaboratore così come un relativo prodotto informatico, caricabile nella memoria di almeno un elaboratore e comprendente porzioni di codice del software atte ad eseguire le fasi del procedimento dell'invenzione quando il prodotto informatico viene eseguito su un elaboratore. Nel presente contesto, il riferimento ad un tale prodotto informatico è inteso essere equivalente al riferimento ad un mezzo leggibile da elaboratore contenente istruzioni per controllare un sistema computerizzato per coordinare il funzionamento del procedimento dell'invenzione. Il riferimento ad "almeno un elaboratore" è chiaramente volto a sottolineare la possibilità per la presente invenzione di essere implementata in maniera distribuita/modulare.

Le rivendicazioni annesse sono parte integrante della descrizione dell'invenzione.

In varie forme di realizzazione, i biosegnali ricevuti da un dispositivo di monitoraggio remoto della salute possono essere utilizzati per taggare un contenuto multimediale con una serie aggiuntiva di metadati

personali.

In varie forme di realizzazione, tali metadati personali possono trasmettere informazioni, eventualmente confidenziali, relative allo stato emotivo e di salute personali. In varie forme di realizzazione, che devono trattare informazioni "sensibili", i dati personali possono essere codificati e una chiave corrispondente può essere distribuita a un gruppo limitato di persone di fiducia.

In varie forme di realizzazione, un codificatore può essere collegato ad uno specifico dispositivo di monitoraggio remoto della salute, che ottiene informazioni da una persona specifica. In varie forme di realizzazione, l'output di un tale codificatore può essere un contenuto multimediale convenzionale con una serie aggiuntiva di metadati unita al collettore del file di contenuto.

In varie forme di realizzazione, una funzione può estrarre una tag dallo stato di salute/emotivo come significato semantico che può essere percepito da un utente finale (ad esempio "triste", "arrabbiato", "in gran forma", ecc.).

Varie forme di realizzazione possono essere basate sul riconoscimento che una delle principali sfide dei sistemi di reperimento semantico è quella di sfruttare l'interazione uomo-macchina attraverso cui gli utenti finali taggano il proprio contenuto multimediale con etichette e riferimenti semantici. Gli utenti finali non eseguono quest'azione poiché è una perdita di tempo, è noioso e irritante.

A tal proposito, gli inventori hanno notato che, nel campo dell'ingegneria biomedica sono contemplate nuove soluzioni che daranno una svolta all'ambito del monitoraggio remoto della salute quali ad esempio

dispositivi body-gateway indossabili e comodi in grado di misurare gli indici biologici e fisiologici. Una volta parte della vita ordinaria delle persone (specialmente in una società senescente), lo sfruttamento degli indici resi disponibili da tali dispositivi può essere preso in considerazione anche per altre attività quotidiane.

Ad esempio, potrebbe essere possibile sfruttare informazioni derivate dall'analisi vocale, dall'analisi del respiro, dall'analisi dell'elettrocardiogramma, dall'analisi del passo, dall'analisi dell'altezza, dall'analisi della pressione sanguigna, e altre, al fine di creare etichette con lo scopo di taggare contenuti multimediali (MM).

Varie forme di realizzazione mirano a rilevare e dedurre lo stato delle persone acquisendo al contempo contenuto multimediale aggiuntivo: l'output può essere un tipo aggiuntivo di tag che è unita allo stesso contenuto.

Varie forme di realizzazione possono essere messe in relazione con l'applicazione del reperimento semantico mediante interrogazione di testo con parola chiave. In varie forme di realizzazione un codificatore può infatti produrre solamente una tag semantica aggiuntiva.

Varie forme di realizzazione possono essere incentrate su un modulo in un'architettura completa, allo scopo di analizzare e processare bio-indici per rilevare uno stato emotivo/fisiologico. In varie forme di realizzazione, l'elaborazione di tali bio-segnali può portare a tag più sofisticate in alternativa a puri campioni del valore del bio-segnale nel tempo.

In varie forme di realizzazione, un tale modulo può essere in grado di rilevare stati emotivi quali la paura, l'ansietà e il rilassamento; variazioni in tale/i stato/i

possono essere rilevate e correlate a eventi che innescano tali variazioni.

Ciò può accadere specialmente in collegamento con eventi in cui una variazione improvvisa nell'ambiente e nelle condizioni esterne ha portato ad una variazione nel profilo del parametro morfologico della persona, ad esempio, solamente a titolo esemplificativo, il battito cardiaco.

In varie forme di realizzazione, un'architettura di elaborazione può trasmettere informazioni provenienti dal dispositivo di monitoraggio remoto della salute; successivamente, il tipo di tag prodotte può essere processato a diversi livelli di astrazione semantica sebbene si faccia affidamento alla stessa architettura di sistema; infine, in alcune forme di realizzazione un codificatore può unire una tag ad un contenuto multimediale, fornendo con tale tag informazioni personali di una persona specifica che è, in generale, il titolare del contenuto foto/video o il titolare del dispositivo che ha prodotto il contenuto MM.

Gli inventori hanno notato che le implicazioni sociali e l'impatto di certe forme di realizzazione possono essere molto forti specialmente se in relazione col profilo "aggregato" dell'intera comunità. Ad esempio, eventi quali un gol segnato in uno stadio, un evento in TV, una canzone ascoltata ad un concerto, un fulmine nel cielo e altri eventi vissuti, possono essere associati a contenuti multimediali eventualmente in grado di trasmettere informazioni aggiuntive riguardo al soggetto rappresentato in una fotografia o in un video, perseguendo pertanto l'obiettivo di far sentire tale soggetto "connesso".

Varie forme di realizzazione pertanto permettono di

sfruttare e collegare contenuti multimediali ad informazioni più personali relative al "titolare" di tali contenuti multimediali.

In varie forme di realizzazione ciò può risultare dal fatto di mescolare i contributi provenienti dai dispositivi "body gateway" per il monitoraggio della salute e dai multimedia.

Varie forme di realizzazione possono riguardare un'architettura di sistema che interfaccia un generatore di contenuti multimediali con un RHMS (soluzione di monitoraggio remoto della salute), allo scopo di estrarre, processare, e successivamente unire al collettore di un contenuto multimediale un'ampia serie di tag aggiuntive che comprendono, senza per questo esserne limitate, il biostato, lo stato emotivo, il GPS, l'etichetta di testo, l'ID utente.

#### Breve descrizione delle figure

Varie forme di realizzazione saranno ora descritte a titolo esemplificativo, con riferimento alle figure allegate, in cui:

- la figura 1, comprendente due porzioni designate rispettivamente con 1a e 1b, rappresenta varie fasi nelle forme di realizzazione;

- la figura 2 rappresenta una forma di realizzazione di un codificatore semantico;

- la figura 3 mostra diversi tipi di connessioni in varie forme di realizzazione;

- le figure 4 e 5 rappresentano forme di realizzazione rispettivamente di un codificatore e di un decodificatore;

- le figure 6 e 7, con la figura 7 comprendente due porzioni designate rispettivamente con 7a e 7b, rappresentano le fasi di codifica e decodifica nelle forme

di realizzazione;

- le figure da 8 a 10 rappresentano varie fasi nelle forme di realizzazione.

#### Descrizione dettagliata

Nella seguente descrizione, vengono forniti numerosi dettagli specifici per una comprensione approfondita delle forme di realizzazione. Le forme di realizzazione possono essere eseguite senza uno o più dettagli specifici, o con altri metodi, componenti, materiali, ecc. In altri casi, strutture, materiali o operazioni noti non sono mostrati o descritti in dettaglio per evitare di rendere oscuri i vari aspetti delle forme di realizzazione.

Il riferimento ad "una forma di realizzazione" nell'ambito di questa descrizione sta ad indicare che una particolare configurazione, struttura o caratteristica descritta in relazione alla forma di realizzazione è compresa in almeno una forma di realizzazione. Quindi, frasi come "in una forma di realizzazione", eventualmente presenti in diversi luoghi di questa descrizione non sono necessariamente riferite alla stessa forma di realizzazione. Inoltre, particolari conformazioni, strutture o caratteristiche possono essere combinate in ogni modo adeguato in una o più forme di realizzazione.

I riferimenti qui utilizzati sono soltanto per comodità e non definiscono dunque l'ambito di tutela o la portata delle forme di realizzazione.

Questa descrizione si riferisce, a titolo esemplificativo, ad un sistema di architettura di codificatore realizzato per un procedimento di taggatura automatica di metadati da utilizzare in associazione con camere digitali.

Nel diagramma di flusso delle figure, i blocchi possono essere o dati informativi, un elemento di informazione, un elemento di metadati, una stringa di dati, oppure i blocchi del modulo logico quale un motore, un blocco logico, un algoritmo, e così via.

Le due porzioni della figura 1, designate con 1a e 1b, rispettivamente, mettono a confronto un approccio convenzionale (figura 1a) e un approccio qui proposto per la prima volta (figura 1b) per produrre contenuti multimediali a partire da dati e metadati grezzi. In alcune forme di realizzazione, tale risultato può essere raggiunto utilizzando un codificatore.

Con riferimento alla figura 1a, l'output del codificatore può comprendere un blocco di tag exif 10 (dove exif sta per formato file immagine scambiabile), un blocco di metadati 20, e un'immagine (dati grezzi) 30; le tag exif sono una specifica per il formato di file immagine utilizzato dalla maggior parte dei marchi di camere digitali. Il contenuto di metadati può comprendere ad esempio informazioni sulla dimensione e sul nome del file.

Nel caso della figura 1b un tipo di formato di metadati è prodotto comprendendo metadati aggiuntivi, anche indicati come metadati personali.

In alcune forme di realizzazione, l'output del codificatore nella figura 1b può anche comprendere un blocco di metadati personali 40. In varie forme di realizzazione, i metadati personali 40 possono essere criptati.

In varie forme di realizzazione, il codificatore della figura 1b può affidarsi ad un'intera architettura che può produrre anche metadati personali come descritto nel seguito.

In varie forme di realizzazione, l'architettura del codificatore può taggare il contenuto multimediale con le informazioni sullo stato di salute/emotivo. In varie forme di realizzazione, l'architettura del codificatore è in grado di gestire tag semantiche multiple al fine di accrescere la capacità e l'efficienza di un sistema di reperimento semantico e aumentare l'esperienza dell'utente in uno scenario di social network.

Un'architettura esemplificativa che rappresenta un possibile flusso di segnali è rappresentata in figura 2. Qui, si suppone che un utente U sia dotato di un dispositivo monitor indossabile W che misura gli indici biologici e fisiologici dell'utente U. Il dispositivo monitor indossabile W produce una serie di bio-indici 60. Tali indici 60 vengono resi disponibili come input per una funzione "sentire" 70 e una funzione "salute" 80. L'output della funzione sentire 70 è l'informazione "stato emotivo" 90 e l'output della funzione salute 80 è informazione "stato di salute" 100.

Il numero di riferimento 50 indica globalmente l'architettura semantica esemplificativa del codificatore.

Qui, le camere digitali o videocamere digitali DC possono produrre tag exif 10 per ciascuna fotografia 5.

In alcune forme di realizzazione un modulo GPS può produrre informazioni di locazione 140. Per ciascun utente U un blocco 150 può produrre contenuto ID relativo a quello specifico utente U. Anche informazioni vocali 160 possono essere utilizzate come parametro di input per la funzione sentire 70.

Un blocco 110 può comprendere ad esempio una lista di categorie di salute, un avatar e un ritratto del titolare, mentre un blocco 120 può comprendere una lista di categorie

emotive in aggiunta all'avatar e al ritratto del titolare. Un'etichetta di testo aggiuntiva 170 può essere aggiunta.

Tutti questi elementi di informazione possono essere mescolati in un blocco 180 allo scopo di produrre nuove tag aggiuntive 130 da collegare ad esempio ad una fotografia 30.

La figura 3 è esemplificativa di una soluzione di monitoraggio remoto della salute (RHMS) comunicante con le camere digitali DC mediante una connessione.

Qui, gli utenti U dotati di dispositivi di monitoraggio indossabili W possono essere connessi alle camere digitali DC in diversi modi. Tale connessione ha lo scopo di accoppiare due dispositivi (W e DC) e può ad esempio essere con fili/senza fili (wired/wireless); la connessione può essere raggiunta mediante mezzi noti (ad esempio, mediante tipi specifici di collegamenti radio), pertanto risulta superfluo fornirne una descrizione dettagliata.

Esempi di connessioni possono essere:

- connessione a breve portata 200 (Wireless, Bluetooth, Infrarossi, Frequenza radio), e
- connessioni a lunga portata 210 mediante qualunque tipo di standard wireless per la comunicazione di dati o vocale (GPS, UMTS, EDGE, proprio come accade con Internet attraverso due indirizzi IP 220 e 230).

In varie forme di realizzazione, due dispositivi generici possono essere provvisti di indirizzi IP. Perciò, un generatore di contenuti multimediali e un dispositivo di monitoraggio remoto della salute possono essere collegati una volta provvisti di indirizzi IP a partire da due posti qualunque nel mondo.

In varie forme di realizzazione, il meccanismo di

accoppiamento qui considerato può comprendere la possibilità di accoppiare un contenuto multimediale non necessariamente ad un produttore di contenuto multimediale, ma a qualunque dispositivo possibile che fornisca altri metadati, ossia bio-metadati.

Pertanto, in varie forme di realizzazione, un codificatore semantico può ricevere contenuto multimediale come input da un generatore di contenuto multimediale, insieme a bio-segnali relativi ad una persona localizzata altrove, anche se non direttamente collegata ad uno stesso evento sociale.

Le figure 4 e 5 rappresentano esempi di gestione di metadati (codificatore vs decodificatore).

Nella forma di realizzazione esemplificativa considerata, tutte le tag sono gestite sul lato del codificatore, si veda figura 4.

In questo esempio, i metadati "tradizionali" 300, relativi ad esempio ad una fotografia 30', possono comprendere dati GPS (o informazioni di locazione) 140, un'etichetta di testo 170, tag exif 10 e metadati di base 20 (quali dimensione, nome file, ecc.).

In questo esempio, i "nuovi" metadati aggiuntivi 310 possono comprendere contenuto ID 150, bio-dati 60 e informazioni sul tono vocale 160.

In varie forme di realizzazione, l'ultima informazione dei metadati aggiuntivi 310 può essere processata da un blocco di elaborazione semantica 240 e poi codificata mediante un blocco di codifica 250 allo scopo di produrre metadati personali 40.

In tal modo, le informazioni collegate alla fotografia 30 sono arricchite con metadati personali 40.

Uno schema duale esemplificativo sul lato del

decodificatore è mostrato in figura 5. Qui, i metadati personali 40 sono decodificati mediante un blocco di decodifica 260 e i risultati sono analizzati in un blocco di analisi semantica 245.

Il blocco del processore semantico 240 corrisponde in maniera logica all'analisi semantica nel decodificatore.

In varie forme di realizzazione, le tag che l'utente finale è in grado di leggere possono essere utilizzate mediante un'interrogazione testo/tag per scopi di reperimento.

Ad esempio, l'input dei bio-dati al lato del codificatore può diventare, ad esempio un elemento di informazione sullo stato di salute al lato del decodificatore, e tali informazioni sullo stato di salute possono diventare metadati che sono utilizzati per taggare il contenuto; queste possono allora essere lette o da un utente finale o essere utilizzate per interrogazioni in sessioni di ricerca di contenuto multimediale.

In varie forme di realizzazione, le informazioni di metadati che sono considerate confidenziali dal titolare possono essere, in maniera opzionale, nascoste mediante chiavi di codifica che sono distribuite solamente dal titolare di tali informazioni.

In varie forme di realizzazione, le informazioni confidenziali possono essere associate ad un produttore ID che non è necessariamente il titolare del contenuto, o il titolare della camera che ha scattato/ripreso il contenuto multimediale.

In varie forme di realizzazione, l'ID in questione può essere un ID che collega le informazioni provenienti dalla soluzione di monitoraggio remoto della salute che può essere accoppiata alla camera mediante ad esempio una

connessione di rete (di qualunque tipo noto). Ad esempio, l'accoppiamento può avvenire mediante dispositivi che si trovano a breve portata "Bluetooth", o, in maniera opzionale, l'accoppiamento può avvenire con qualunque dispositivo raggiungibile di soluzione di monitoraggio remoto della salute.

In varie forme di realizzazione, può essere possibile unire i metadati provenienti dalla soluzione di monitoraggio remoto della salute di un caro amico negli USA mentre scatta una fotografia in Europa.

In varie forme di realizzazione, l'architettura esemplificativa considerata può essere pienamente determinata dall'accesso alle informazioni disponibili mediante il dispositivo gateway portatile di monitoraggio della salute.

In varie forme di realizzazione, i metadati che si riferiscono alle informazioni che l'utente vuole nascondere possono richiedere un meccanismo di codifica dei dati e di una firma digitale. In varie forme di realizzazione, mentre il meccanismo di codifica può nascondere il valore vero e proprio delle tag, la firma digitale può assicurare che tali valori siano associati ad un produttore ID. In quel caso, il produttore ID può riferirsi al produttore dei bio-segnali collegati alla soluzione di monitoraggio remoto della salute.

La figura 6 rappresenta in maniera schematica un tipo di processo di firma digitale "dell'autore".

In una tale forma di realizzazione esemplificativa, un vettore di informazioni comprendente ad esempio contenuto ID 150, bio-indice 60, stato di salute 100, stato emotivo 90, geo-tag o informazioni di locazione 140, ed etichetta di testo 140 può essere reso disponibile come input ad un

blocco hash 400. Allora l'output del blocco hash 400 può essere inviato ad un blocco digest 405, e successivamente ad una fase di codifica 250. La fase di codifica 250 può anche ricevere una chiave di codifica 410 allo scopo di produrre informazione codificata hash 415 e una firma digitale 420.

I processi di codifica e decodifica esemplificativi ai lati del produttore e del consumatore sono illustrati rispettivamente in figura 7a e in figura 7b.

In tali forme di realizzazione esemplificative, i metadati personali 40 possono essere codificati in una fase 250 utilizzando una chiave privata 430 allo scopo di ottenere metadati personali codificati 40'. Al lato del produttore i metadati codificati 40' possono essere decodificati in una fase 260 utilizzando una chiave pubblica 435 per ottenere i metadati originali 40.

Mentre la chiave privata 430 può essere utilizzata per codificare (nascondere) i metadati personali al lato del produttore, una chiave pubblica 435 può essere utilizzata per decodificare i metadati personali al lato del consumatore. La chiave pubblica 435 può essere distribuita da un canale segreto solamente alle persone di fiducia del produttore.

In varie forme di realizzazione, una funzione esemplificativa di bio-stato  $B(t)$  può rappresentare lo stato di un'intera serie di bio-segnali  $b_1(t)$ ,  $b_2(t)$ ..  $b_i(t)$ , ...  $b_N(t)$ , dove  $N$  è il numero dei segnali rilevati dal sistema di soluzione di monitoraggio remoto della salute; la funzione  $B(t)$  di bio-stato può essere ad esempio una funzione vettore i cui elementi sono rappresentati da funzioni  $N$  scalari.

A tal proposito, la figura 8 rappresenta

schematicamente una catena di fasi di computazione eseguite sui bio-segnali 60. La catena esemplificativa qui considerata descrive un modo possibile di definire bio-tag e di associare tali bio-tag ad un contenuto multimediale.

Infatti, a causa di vincoli in termini di bassa energia, e, eventualmente, a causa di una rilevanza semantica dei bio-segnali ad un evento, in varie forme di realizzazione, una funzione  $B(t)$  di bio-stato dell'utente di RMHS può essere campionata ad intervalli di tempo  $\Delta T$ ; dopodiché la funzione  $B(t)$  di bio-stato può essere unita al contenuto multimediale generato dal generatore di contenuto multimediale accoppiato al dispositivo di RMHS generato.

In varie forme di realizzazione, una corrispondenza naturale e intuitiva può basarsi sul principio di coerenza temporale, ossia il contenuto multimediale generato nel tempo  $t_i$  può essere associato al bio-stato nello stesso tempo  $t_i$ .

In varie forme di realizzazione, l'associazione tra la generazione del contenuto e il bio-stato può essere resa tempo-coerente, pertanto non è necessario che il  $\Delta T$  sia troppo breve (microsecondi) o troppo lungo (ore).

Eventualmente, in varie forme di realizzazione, un valore per  $\Delta T$  può essere (pre)stabilito ad un valore standard dalla configurazione di avvio del dispositivo o mediante configurazione manuale dell'utente stesso.

Vi è anche la possibilità di differenziare l'intervallo temporale  $\Delta T$  per ogni bio-indice 60 e lungo la linea temporale.

La figura 9 illustra una forma di realizzazione esemplificativa dove l'intervallo di campionatura è una funzione del tempo. Lo scopo può essere quello di ottimizzare l'utilizzo delle risorse quando la bio-attività

è costante per un certo periodo di tempo, o quando diventa molto variabile.

Questa caratteristica può essere espressa dall'introduzione di una funzione vettore  $\Delta T$ , come indicato nell'equazione 1.

$$\Delta T = (f_{\Delta T1}(t), f_{\Delta T2}(t), \dots, f_{\Delta Ti}(t), \dots, f_{\Delta TN}(t)) \text{ Eq. 1}$$

In alcune forme di realizzazione, vi è la possibilità di memorizzare l'intera bio-storia dell'RMHS per ciascun bio-indice in una memoria di archiviazione del bio-stato (MBS). In quel caso, un buffer di valori memorizzati nella matrice dei valori di bio-indice (MBS) può essere reso disponibile per essere interrogato, in caso di bisogno, quando il contenuto multimediale viene generato. Inoltre, in varie forme di realizzazione, vi è la possibilità di comprimere la storia in qualunque formato compresso e accedere alla memoria di archiviazione (MBS) con qualunque tipo di algoritmo nel dominio compresso e reperire i relativi bio-indici da utilizzare come metadati per il contenuto multimediale generato, o altre applicazioni di strato superiore.

Inoltre, in alcune forme di realizzazione, vi è la possibilità di associare il contenuto multimediale al bio-stato trascurando una corrispondenza temporale troppo ovvia. Il senso comune suggerisce, infatti, che se un contenuto è generato in un tempo generico  $t$ , allora l'associazione può essere basata su una corrispondenza temporale con il bio-stato campionato nel tempo  $t_0$  in cui  $t$  è:  $t_0 \leq t < t_1$ . Tuttavia, in alcune forme di realizzazione è possibile far corrispondere il contenuto generato nel tempo  $t_x$  con il bio-stato campionato nel  $t_i$  in cui  $x \ll i$  o

i << x.

Pertanto è possibile unire il bio-stato ad un contenuto mantenendo una distanza temporale dalla generazione del contenuto multimediale.

In varie forme di realizzazione, i bio-dati possono essere campionati e il bio-stato unito al contenuto multimediale. In varie forme di realizzazione, le tag aggiuntive possono essere generate a partire da bio-indici e poste nel codificatore semantico per produrre tag più sofisticate per essere utilizzate dall'architettura del sistema come metadati per il contenuto multimediale.

Verrà ora descritto un processo esemplificativo di estrazione delle informazioni di bio-stato a partire da segnali fisiologici.

A tal proposito, gli inventori hanno notato che il corpo umano è un'eccellente fonte di informazioni. Registrando e analizzando diversi segnali fisiologici, è possibile valutare dati preziosi relativi allo stato di salute dell'utente. Ciascun segnale trasmette un diverso tipo di informazione, perché è basato su un fenomeno fisiologico differente:

l'ECG registra l'attività elettrica del cuore, il PPG registra otticamente la perfusione sanguigna, un termometro può registrare le variazioni nella temperatura della pelle, un microfono può registrare il segnale audio del battito cardiaco.

Processando tali segnali, si può ottenere uno specifico indice fisiologico, ad esempio come "valore sintetico" ad esempio la frequenza cardiaca, la frequenza respiratoria, la temperatura, ecc. Combinando tutti questi diversi indici fisiologici, si può valutare lo stato corrente dell'utente. Un'alta frequenza cardiaca indica che

l'utente si è affaticato probabilmente per una lunga camminata, un'elevata temperatura corporea indica che è stata una giornata veramente calda, informazioni queste che possono accrescere enormemente la completezza dei dati codificati con il file multimediale. Infine tali indici fisiologici possono essere utilizzati per stimare le emozioni della persona, permettendo di comprendere ad esempio se l'utente è elettrizzato, arrabbiato, assonnato e così via. La figura 10 rappresenta un modo esemplificativo di estrazione in una fase 500 di bio-dati a partire da bio-segnali 60, che sono segnali analogici, per generare successivamente tag 505. Tali tag 505 possono essere inviate come input al codificatore semantico 50. Ad esempio,  $tag_i = tag(t)$ ,  $tag_{i+1} = tag(t+\Delta)$ ,  $tag_{i+2} = tag(t+2\Delta)$ . L'approccio adottato nel codificatore di questa forma di realizzazione esemplificativa può essere rappresentato dalle seguenti caratteristiche.

- Stato di salute (modulo di processo): i bio-segnali sono processati per reperire o inferire lo stato di salute dell'utente finale mediante una funzione  $F_{health}$  (Eq. 2). La funzione  $F_{health}$  può operare come segue:

$$F_{health} = f(B_1, B_2, \dots B_i, \dots B_n) \quad \text{Eq. 2}$$

dove  $B_i$  rappresenta qualunque bio-indice che può essere disponibile dal rilevamento dell'RHMS o dello stato d'animo a partire dalla voce. L'output della funzione  $F_{health}$  può comprendere:

a) l'utilizzo di una serie di stati di salute predefiniti come icone, loghi, sfondi colorati di diverso tipo o saturazione/manipolazione del colore per la stessa fotografia (ad esempio, il ritratto del titolare del contenuto), che mirano a rappresentare un numero limitato di stati di salute;

b) l'utilizzo della rappresentazione di cui sopra per indicare al visore dell'utente finale quale era lo stato di salute al momento dell'acquisizione del contenuto multimediale. - Stato emotivo (modulo di processo): la correlazione dei bio-segnali con un motore di riconoscimento dello stato d'animo a partire dalla voce per reperire i segnali analogici disponibili mediante qualunque tipo di connessione wired/wifi (solitamente Bluetooth) per inferire lo stato emotivo dell'utente finale attraverso una funzione  $F_{feel}$  (Eq. 3).

$$F_{feel} = f(B_1, B_2, \dots B_i, \dots B_n) \quad \text{Eq. 3.}$$

L'output della funzione  $F_{feel}$  può comprendere:

a) l'utilizzo di una serie di stati emotivi predefiniti come icone, loghi, sfondi colorati di diverso tipo o saturazione/manipolazione del colore per la stessa fotografia, che mirano a rappresentare un numero limitato di stati emotivi;

b) l'utilizzo della rappresentazione di cui sopra per indicare al visore dell'utente finale quale era lo stato emotivo al momento dell'acquisizione del contenuto multimediale.

- Per includere lo stato di salute come tag semantica nel contenuto multimediale.

- Per includere lo stato emotivo come tag semantica nel contenuto multimediale.

- Per includere le GeoTag (il luogo al momento della acquisizione).

- Per includere l'identificazione del titolare del contenuto del contenuto multimediale mediante qualunque tipo di informazione che sia univocamente collegata al titolare. Tale forma di identificazione comprende qualunque tipo di tag multimediale o testuale quale un logo,

un'icona, un piccolo ritratto digitale, un file di testo, una tag vocale, una stringa ID, un documento ID personale.

- Per incorporare le GEO tag con i metadati tag exif (uno standard di metadati comuni utilizzato per tutte le camere digitali).

- Per includere etichette di testo specifiche indicate mediante:

a) Una serie predefinita di etichette inserite attraverso un'interfaccia dedicata uomo-macchina;

b) L'impostazione manuale di etichette mediante input della tastiera, o mediante motori automatici di riconoscimento vocale dotati di ampio vocabolario, o mediante altri tipi di interfaccia dedicata uomo-macchina.

- Un meccanismo per proteggere le informazioni sensibili dividendo i metadati tra le Geo tag, le tag exif e le altre tag tradizionali con le informazioni di metadati personali. La geolocalizzazione può, in base alle preferenze dell'utente, essere compresa nella serie di informazioni di metadati personali/confidenziali. Quindi le informazioni confidenziali necessitano di un algoritmo Hash protetto per codificare le tag che l'utente vuole nascondere.

L'architettura proposta può essere rilevabile da qualunque utente finale senza nessuna ingegneria inversa:

- al lato del decodificatore: in qualunque sistema di reperimento poiché i contenuti multimediali possono essere reperiti mediante nuove tag specifiche quali tag bio-indice, stato di salute, stato emotivo. Questo tipo di codificatore può infatti permettere un nuovo tipo di reperimento semantico: reperimento dello stato emotivo/di salute (il contenuto multimediale viene reperito interrogando uno stato emotivo/di salute); e

- al lato del codificatore: gli utenti possono

indossare e connettere alla camera (preferibilmente via wireless) il gateway portatile.

Varie forme di realizzazione possono essere applicate vantaggiosamente alle aree seguenti:

- al lato del decodificatore, tutti i sistemi atti ad eseguire forme di reperimento dell'informazione anche molto semplici, in particolare le "smart" Set Top Boxes, poiché il reperimento viene effettuato tramite tag;

- al lato del codificatore, le camere digitali dove nuove tag possono essere aggiunte ai metadati convenzionali;

- body gateway portatili; e

- relative applicazioni software.

Fatti salvi i principi di base dell'invenzione, i dettagli e le forme di realizzazione possono variare, anche in maniera significativa, rispetto a quanto è stato qui descritto attraverso esempi non limitativi, senza allontanarsi dall'ambito di protezione dell'invenzione, come definito dalle rivendicazioni.

#### RIFERIMENTI BIBLIOGRAFICI

[1] K. H. Kim, S. W. Bang, S. R. Kim, "Emotion recognition system using short-term monitoring of physiological signals", Medical & Biological Engineering & Computing 2004, Vol. 42.

[2] J. Wagner, J. Kim, E. Andrè, "From physiological signals to emotions", Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on.

[3] H.S. Goldstein, R. Eldberg, C.F. Meier, L. Davis, "Relationship of resting blood pressure and heart rate to experienced anger and expressed anger", Psychosomatic Medicine 50:321-329 (1988)

[4] D. Shapiro, D. Jamner, L. Goldstein, I. Delfino, "Striking a chord: Moods, blood pressure, and heart rate in everyday life", Psychophysiology, Volume 38 Issue 2 Pag. 197 - Marzo 2001.

## RIVENDICAZIONI

1. Sistema di taggatura multimediale, comprendente:

- un generatore di contenuti multimediali (DC) per produrre contenuto multimediale (30) taggato con metadati (10, 20), e

- un dispositivo di monitoraggio remoto della salute (W) per misurare e processare (50) una serie di segnali biologici e fisiologici (60) di un utente (U),

in cui detto sistema è configurato per taggare detto contenuto multimediale (30) con tag (505) estratte da una serie di metadati personali (40) ottenuti da detti segnali biologici e fisiologici (60) forniti da detto dispositivo di monitoraggio remoto della salute (W) e contenenti informazioni relative allo stato emotivo e di salute di detto utente (U).

2. Sistema secondo la rivendicazione 1, in cui detta serie di metadati (40) è criptata (250) e unita ad un header di contenuti multimediali.

3. Sistema secondo la rivendicazione 1 o 2, in cui detti segnali biologici e fisiologici (60) sono processati (70, 80) per ottenere informazioni sullo stato emotivo (90) e informazioni sullo stato di salute (100).

4. Sistema secondo la rivendicazione 3, in cui dette informazioni sullo stato emotivo (90) sono selezionate in una lista di categorie emotive (120), e dette informazioni sullo stato di salute (100) sono selezionate in una lista di categorie di salute (110).

5. Sistema secondo una qualunque delle rivendicazioni precedenti da 1 a 4, in cui detto dispositivo di monitoraggio remoto della salute (W) comunica con detto generatore di contenuti multimediali (DC) mediante una connessione a breve portata (200) selezionata tra Wireless, Bluetooth, Infrarossi, Frequenza Radio, o mediante una connessione a lunga portata (210) selezionata tra GPS, UMTS, EDGE.

6. Sistema secondo una qualunque delle rivendicazioni precedenti, in cui il contenuto multimediale generato in un certo istante temporale è associato ai valori di detti metadati personali aggiuntivi (40) misurati nello stesso istante temporale generando una coerenza temporale.

7. Sistema secondo una qualunque delle rivendicazioni precedenti, in cui il sistema è configurato per accoppiare al contenuto multimediale prodotto da una macchina fotografica digitale di un primo utente che scatta una fotografia in una prima posizione i metadati personali (40) relativi ad un secondo utente che indossa un body gateway in una seconda posizione.

8. Sistema secondo una qualunque delle rivendicazioni da 1 a 7, in cui detto sistema mostra lo stato di salute predefinito con icone, loghi, o un diverso sfondo di colori o saturazione/manipolazione del colore.

9. Procedimento implementato mediante elaboratore, comprendente:

- la generazione di contenuti multimediali (30) taggati con metadati (10, 20),

- la misurazione e l'elaborazione (50) di una serie di segnali biologici e fisiologici (60) di un utente (U) ottenuti mediante un dispositivo di monitoraggio remoto della salute (W),

- la taggatura di detto contenuto multimediale (30) con tag (505) estratte da una serie di metadati personali (40) ottenuti da detti segnali biologici e fisiologici (60) forniti da detto dispositivo di monitoraggio remoto della salute (W) e contenenti informazioni relative allo stato emotivo e di salute di detto utente (U).

10. Procedimento implementato mediante elaboratore secondo la rivendicazione 9, come implementato con il sistema di una qualunque delle rivendicazioni da 1 a 8.

11. Prodotto informatico, caricabile nella memoria di almeno un elaboratore e comprendente porzioni di codice del software atte ad eseguire le fasi del procedimento di cui alla rivendicazione 9 o 10 quando il prodotto viene eseguito su un elaboratore.

## CLAIMS

1. A multimedia tagging system, including:  
- a multimedia content generator (DC) for producing multimedia content (30) tagged with metadata (10, 20), and  
- a remote health monitoring device (W) for measuring and processing (50) a set of biological and physiological signals (60) of a user (U),

wherein said system is configured for tagging said multimedia content (30) with tags (505) extracted from a set of personal metadata (40) obtained from said biological and physiological signals (60) provided by said remote health monitoring device (W) and containing information relative to the emotional and health status of said user (U).

2. The system of claim 1, wherein said set of personal metadata (40) is encrypted (250) and attached to the header of the multimedia content.

3. The system of claim 1 or claim 2, wherein said biological and physiological signals (60) are processed (70, 80) to obtain feeling status information (90) and health status information (100).

4. The system of claim 3, wherein said feeling status information (90) is selected among a list of feeling categories (120), and said health status information (100) is selected among a list of health categories (110).

5. The system of any of the previous claims 1 to 4, wherein said remote health monitoring device (W)

communicates with said multimedia content generator (DC) through a short range connection (200) selected among Wireless, Bluetooth, Infra Red, Radio Frequency, or a long range connection (210) selected among GPS, UMTS, EDGE.

6. The system of any of the previous claims, wherein the multimedia content generated at certain instant time is associated with the values of said additional personal metadata (40) measured at the same instant time generating a temporal consistency.

7. The system of any of the previous claims, wherein the system is configured for coupling to the Multimedia content produced by a digital camera of a first user shooting a picture at a first location the personal metadata (40) relative to a second user wearing a body gateway at a second location.

8. The system of any of claims 1 to 7, wherein said system shows the pre-defined health status with icons, logos, or different type background of colors or color saturation/manipulation.

9. A computer-implemented method, including:

- generating multimedia content (30) tagged with metadata (10, 20),

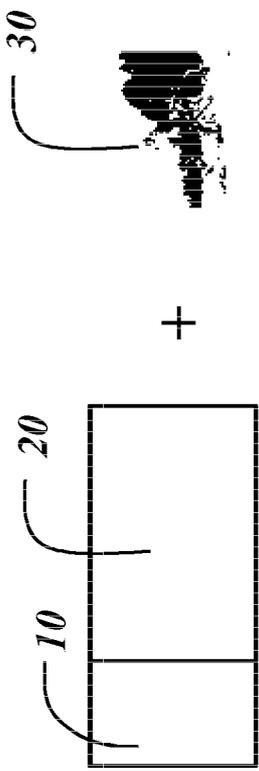
- measuring and processing (50) a set of biological and physiological signals (60) of a user (U) obtained via a remote health monitoring device (W),

- tagging said multimedia content (30) with tags (505) extracted from a set of personal metadata (40) obtained from said biological and physiological signals (60)

provided by said remote health monitoring device (W) and containing information relative to the emotional and health status of said user (U).

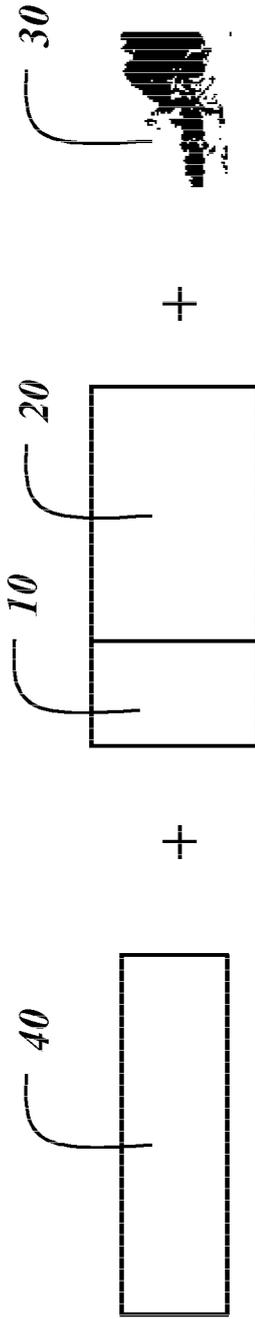
10. The computer-implemented method of claim 9, as implemented with the system of any of claims 1 to 8.

11. A computer program product, loadable in the memory of at least one computer and including software code portions for performing the steps of the method of claim 9 or claim 10 when the product is run on a computer.



*a*

*1/10*



*b*

**FIG. 1**



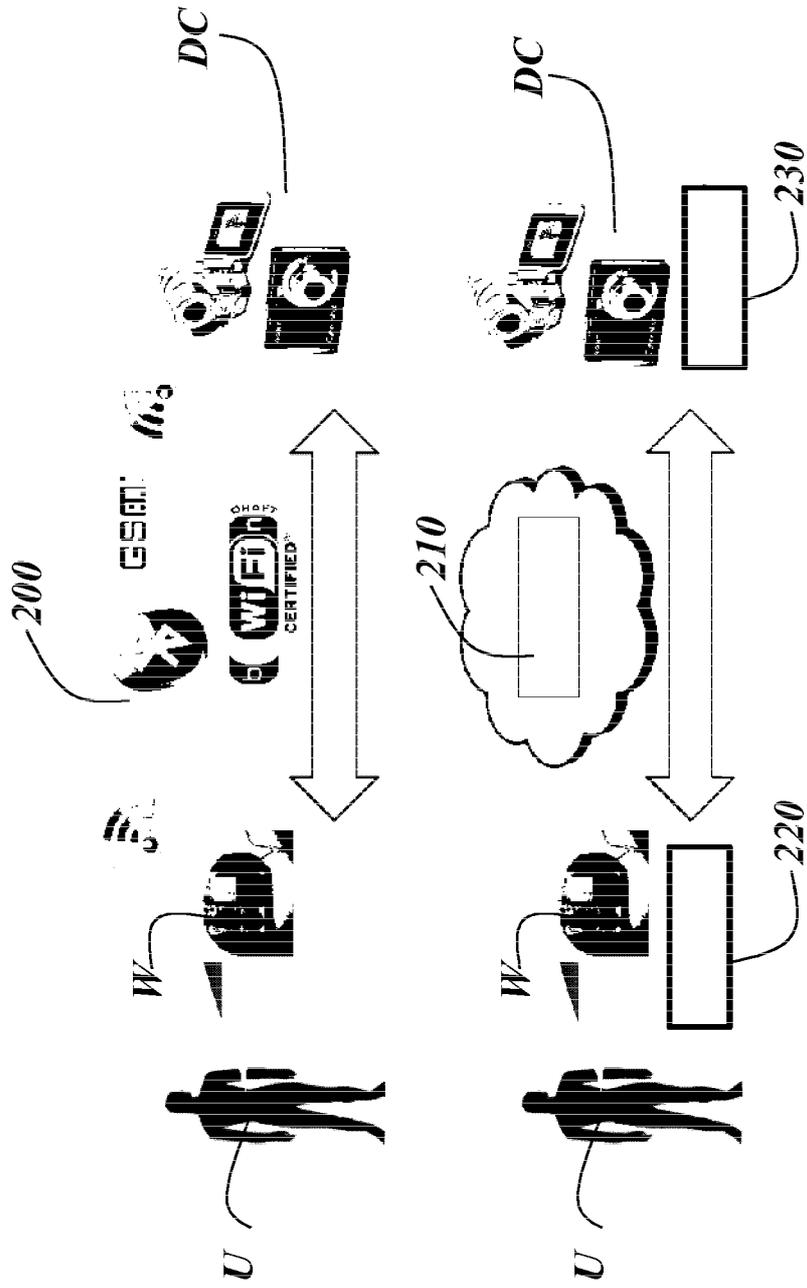


FIG. 3

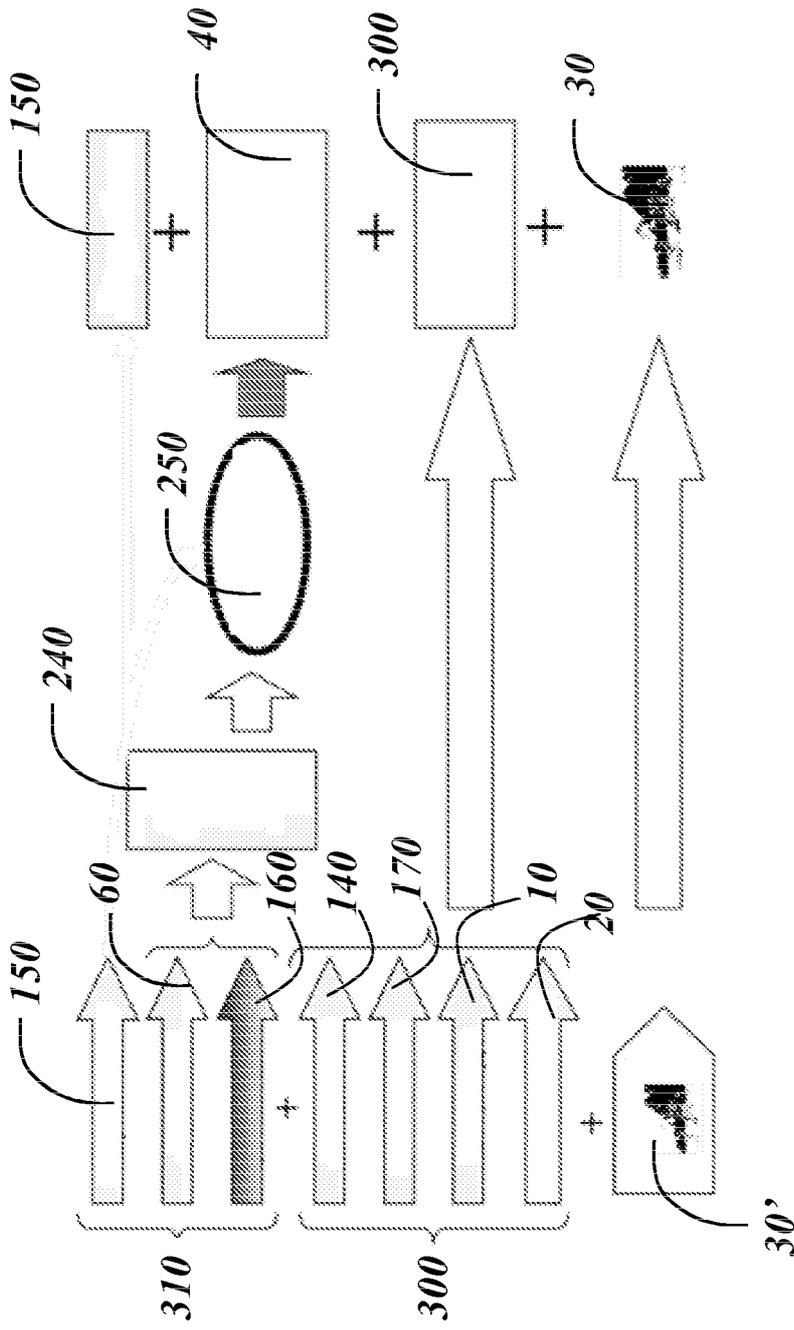


FIG. 4

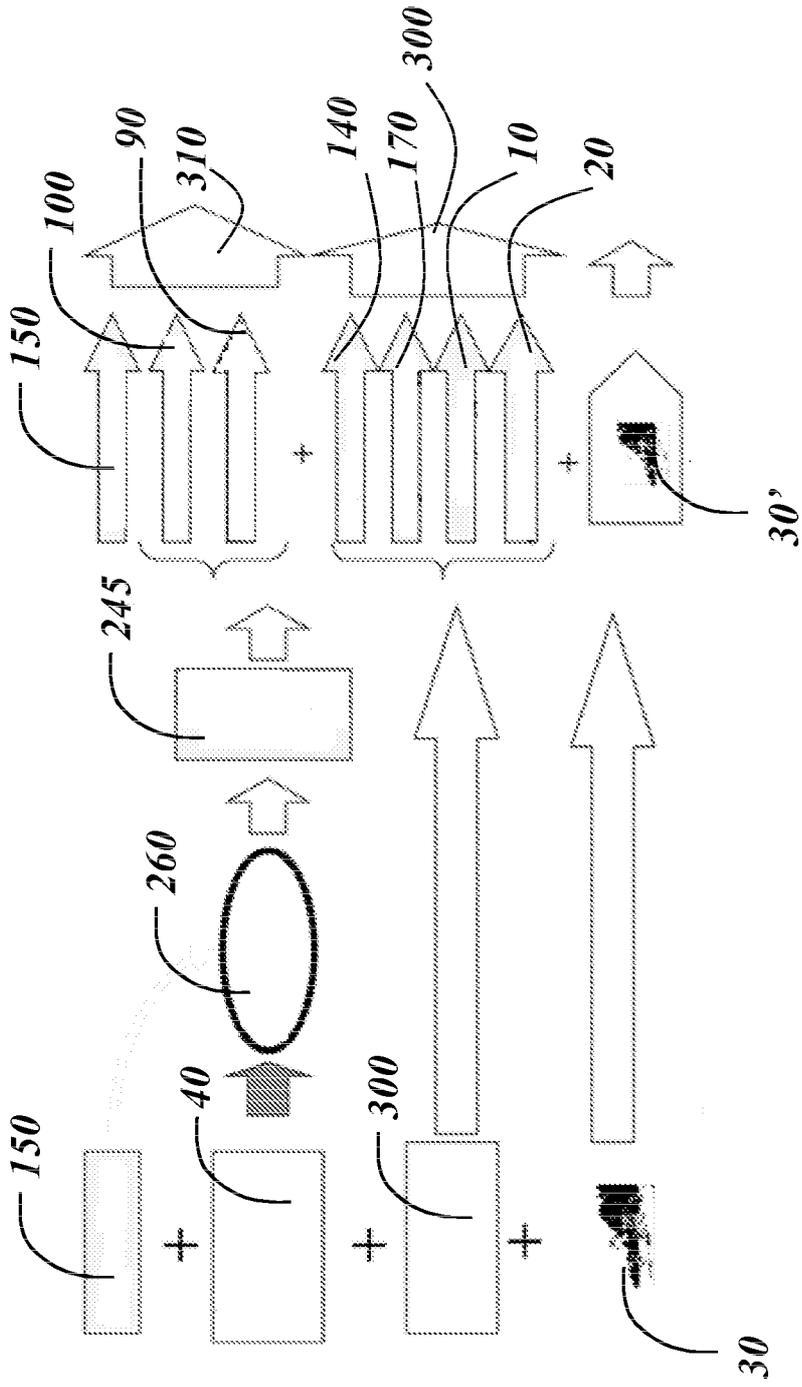


FIG. 5

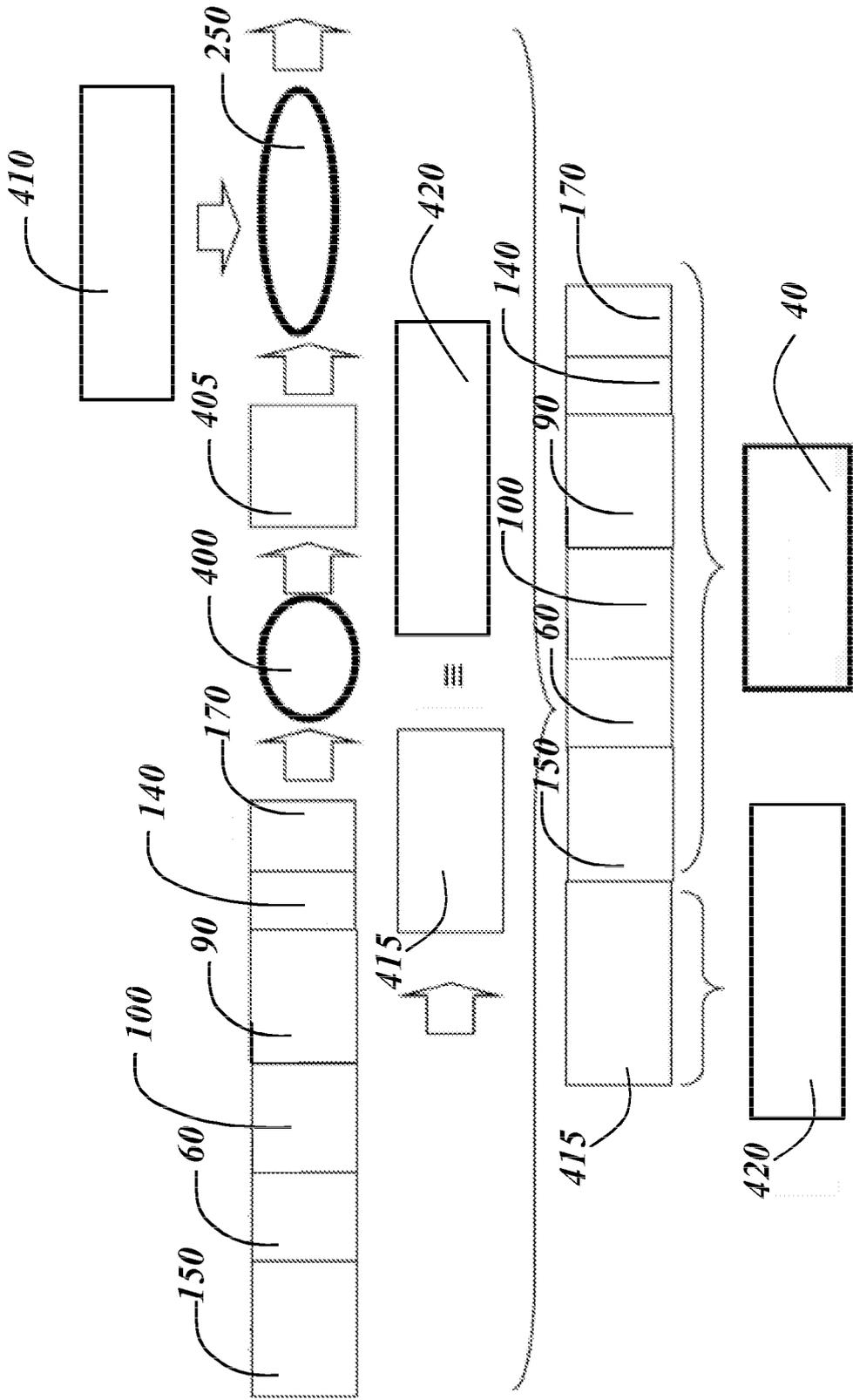


FIG. 6

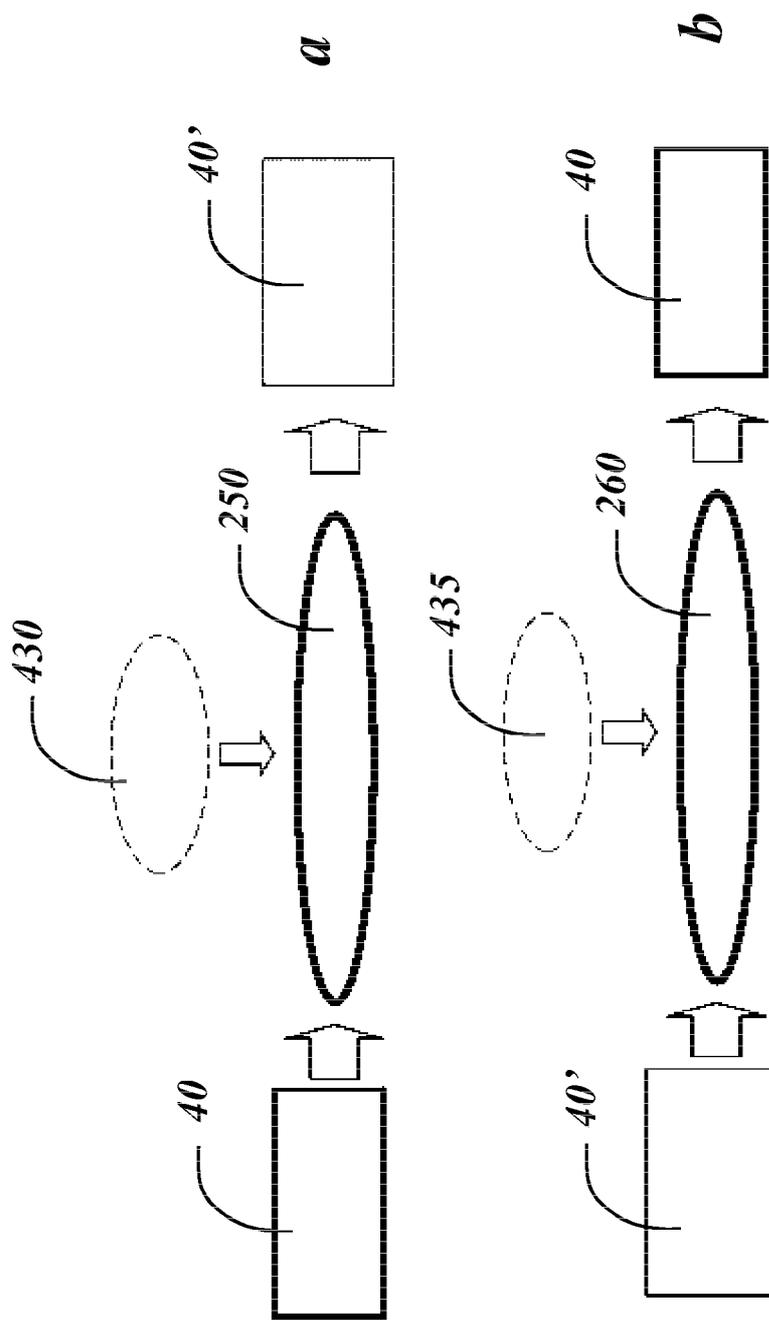


FIG. 7

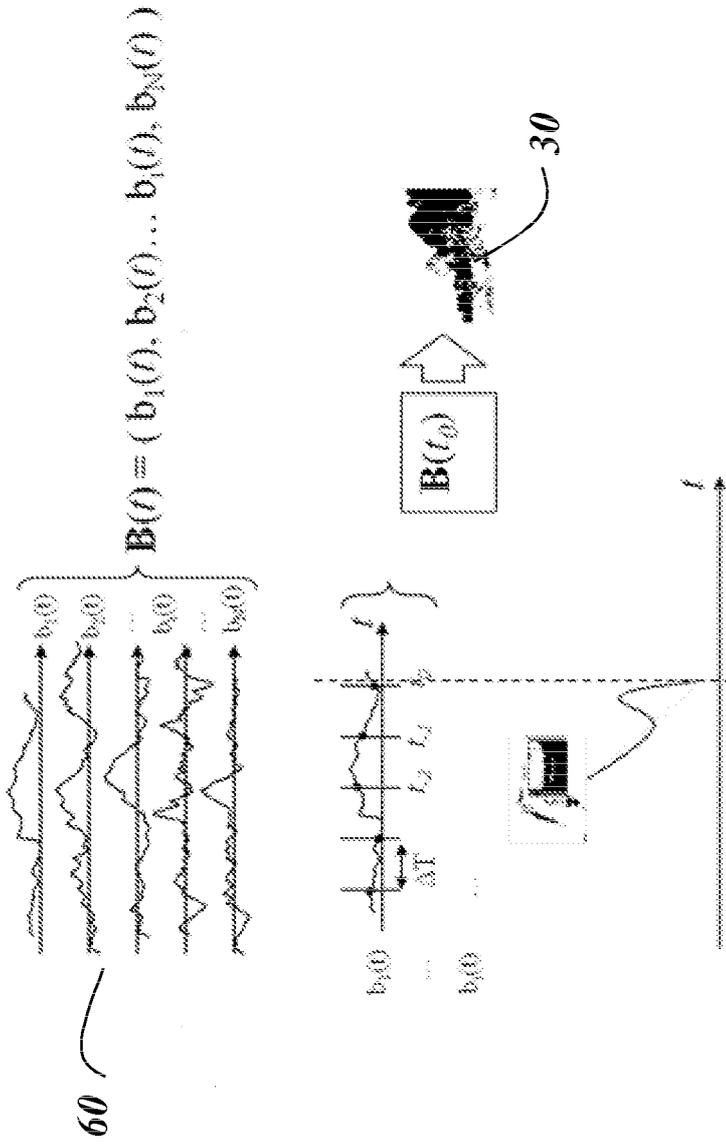
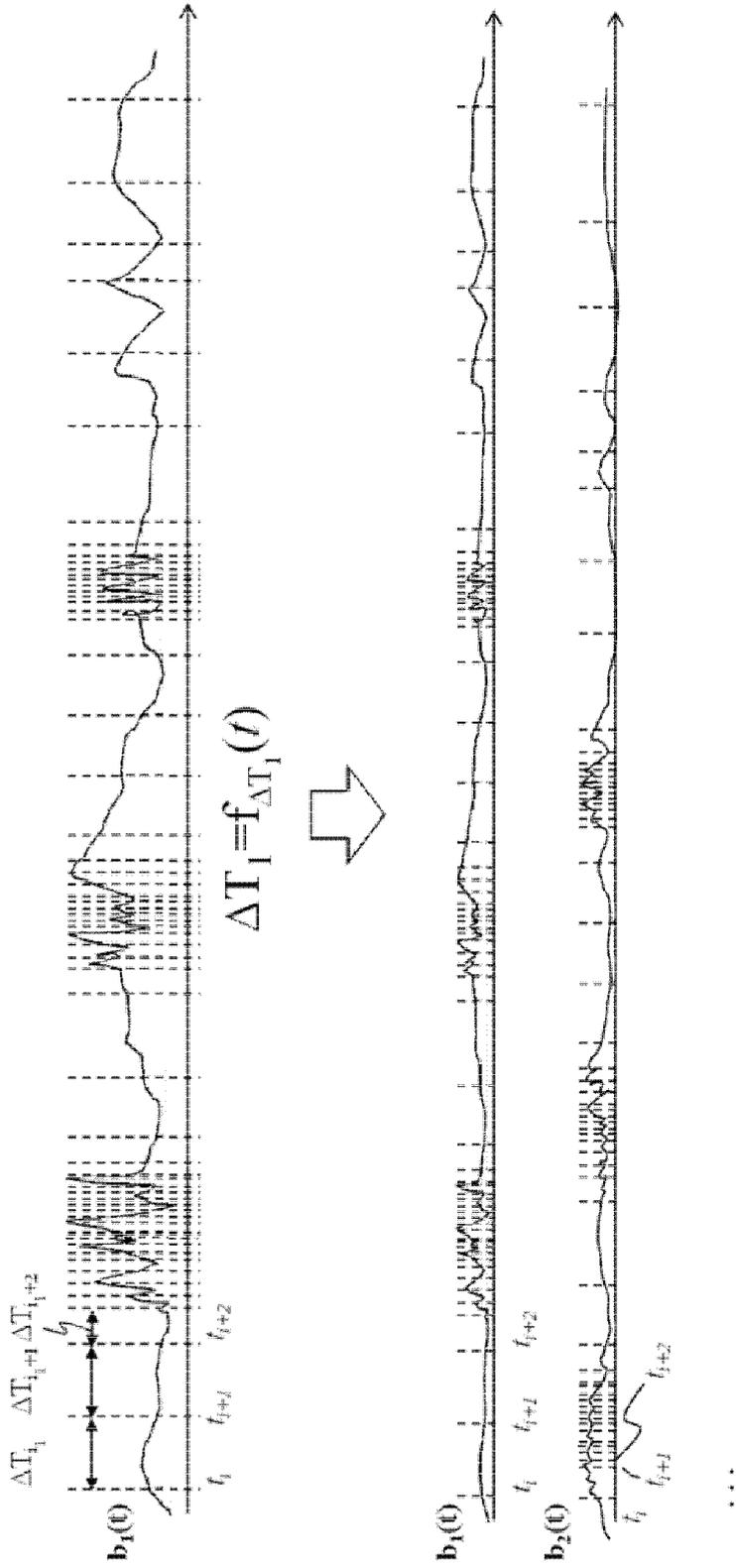


FIG. 8



**FIG. 9**

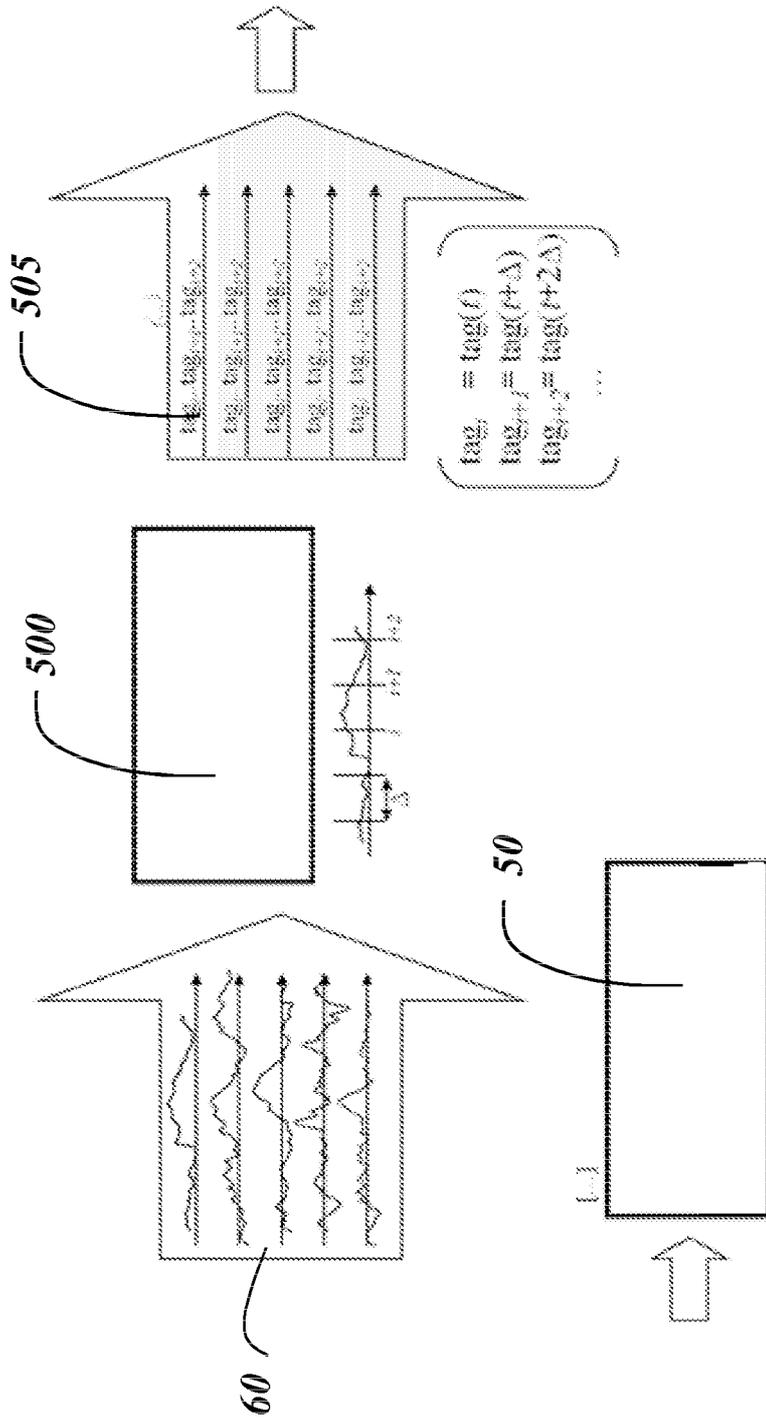


FIG. 10