



MINISTERO DELLO SVILUPPO ECONOMICO  
DIREZIONE GENERALE PER LA LOTTA ALLA CONTRAFFAZIONE  
UFFICIO ITALIANO BREVETTI E MARCHI

DOMANDA NUMERO	102006901471849
Data Deposito	30/11/2006
Data Pubblicazione	30/05/2008

Sezione	Classe	Sottoclasse	Gruppo	Sottogruppo
G	06	K		

Titolo

METODO E APPARATO PER RICONOSCERE TESTO IN UNA IMMAGINE DIGITALE.

P02725/IT/FC

**Titolo:** “Metodo e apparato per riconoscere testo in una immagine digitale.”

### DESCRIZIONE

5           La presente invenzione si riferisce ad un metodo e ad un apparato per riconoscere testo in una immagine digitale, in particolare ad un metodo e ad un apparato in accordo con il preambolo delle rivendicazioni 1 e 27.

          Nella società odierna, la maggior parte delle informazioni sono acquisibili mediante la vista. Nonostante le leggi vigenti stiano incrementando il  
10 numero di informazioni disponibili mediante il tatto (si pensi per esempio ai pulsanti degli ascensori, alle scritte in Braille sui medicinali, alle indicazioni tattili sui pavimenti degli aeroporti o delle stazioni), vi sono ancora moltissimi campi in cui un utente non-vedente o ipo-vedente ha un'estrema difficoltà a fruire delle informazioni disponibili.

15           Per esempio, si pensi alle bevande contenute in bottiglie o in lattine, o al contenuto delle scatolette alimentari, o alle scatole contenenti alimenti, o ai medicinali.

          In tutti questi casi, un utente non-vedente o ipo-vedente può trovarsi con estrema facilità nella situazione di dover ottenere informazioni che sono solo  
20 disponibili tramite stimoli visivi, proprio quelli che l'utente ipo-vedente o non-vedente non può decodificare con facilità.

          Per esempio, tutte le lattine hanno la stessa forma e pressoché le stesse dimensioni, essendo il differente contenuto ampiamente descritto sulla superficie della lattina stessa; le bottiglie di vetro possono contenere  
25 indifferentemente vino, acqua, olio, aceto, liquori, oppure le bottiglie di plastica

con la stessa forma possono contenere trielina o acqua; scatolette con la stessa forma possono contenere fagioli, passata di pomodoro o tonno; le scatole di cartone possono contenere pasta con tempi di cottura estremamente differenti ed, infine, i medicinali, hanno le informazioni relative al nome della medicina ed alla data di scadenza in formato tattile solo sulla confezione: quando il blister  
5 è separato da questa, per un utente non-vedente o ipo-vedente può diventare praticamente impossibile decodificare correttamente queste informazioni partendo solo dalla forma del blister.

Da questi esempi, è chiaro come un errore nella comprensione di queste  
10 informazioni visive può avere conseguenze spiacevoli, dalle lievi conseguenze di mangiare una pasta troppo cotta con un sugo al tonno anziché ai fagioli, fino all'avvelenamento provocato dall'assunzione di un medicinale sbagliato o di un liquido corrosivo al posto di semplice acqua.

Per evitare tali problemi, gli utenti non-vedenti o ipo-vedenti si affidano  
15 generalmente all'aiuto di persone normo-vedenti che possono leggere per loro quelle informazioni importanti, rinunciando magari all'aiuto esterno per informazioni meno rilevanti, come per esempio quelle relative alla marca di un detersivo o al tempo di cottura della pasta.

Sarebbe però desiderabile avere un sistema che permetta agli utenti  
20 ipo-vedenti e/o non-vedenti di svincolarsi dall'aiuto di altre persone, in modo da essere meno dipendenti da costoro e da poter sfruttare anche le informazioni meno importanti, comunque fondamentali per una migliore qualità della vita.

Nella tecnica sono noti dispositivi di acquisizione di immagini digitali, programmi OCR (Optical Character Recognition, ovvero programmi in grado di  
25 convertire il testo contenuto in una immagine digitale in codice testuale) e

sintetizzatori vocali in grado di leggere il risultato dell'OCR.

I dispositivi di acquisizione d'immagine, siano essi scanner, macchine fotografiche (digitali o meno), webcam, scanner a planetario, videocamere digitali o simili risentono della mancanza di precisione nell'impostazione dei parametri di  
5 acquisizione dell'immagine; il risultato è, frequentemente, un'immagine sfuocata e/o mossa, sovra- o sottoesposta, con un'inquadratura errata, con una presenza di riflessi indesiderati, con un'errata taratura del colore, del contrasto e/o della profondità di campo.

A tutti i problemi sopra menzionati relativi alla bontà dell'immagine di  
10 partenza si somma il problema relativo alla forma dell'oggetto fotografato, che, nel caso di una lattina, per esempio, potrebbe richiedere anche ad un utente normo-vedente di ruotare la lattina attorno al suo asse per leggere per intero il nome della marca e quindi conoscerne il contenuto.

A partire da queste immagini è attualmente impossibile per un sistema di  
15 OCR ottenere buoni risultati, in quanto gli OCR attualmente disponibili in commercio sono stati sviluppati sostanzialmente per leggere documenti contenenti testo stampato in bianco su nero, con caratteri regolari e di tipo standard (Arial, Times New Roman, Courier, eccetera), con un corpo di almeno 10-12 punti e allineati lungo la direzione verticale o la direzione orizzontale.

20 I risultati attualmente possibili sono limitati in termini di praticità a, per esempio, la lettura dei foglietti illustrativi dei farmaci.

I problemi sopra indicati di inquadratura, illuminazione e di stabilità dell'immagine sono stati affrontati dal titolare della presente invenzione e risolti almeno parzialmente mediante il dispositivo oggetto del brevetto avente per titolo  
25 "DISPOSITIVO DI ASSISTENZA PER IPOVEDENTI PER SCATTARE

FOTOGRAFIE”, depositato a nome della medesima Richiedente in data 29 novembre 2006.

In vista dello stato della tecnica descritto, scopo della presente invenzione è quello di realizzare un metodo ed un apparato per riconoscere testo in una  
5 immagine digitale che non presenti gli svantaggi dell’arte anteriore.

In accordo con la presente invenzione, tale scopo viene raggiunto mediante un metodo secondo la rivendicazione 1 ed un apparato secondo la rivendicazione 27.

Tale scopo è anche raggiunto un prodotto informatico che, caricato nella  
10 memoria di un computer e fatto lavorare su tale computer, consente di attuare il metodo in accordo con la presente invenzione.

Le caratteristiche ed i vantaggi della presente invenzione risulteranno evidenti dalla seguente descrizione dettagliata di una forma di realizzazione pratica, illustrata a titolo di esempio non limitativo negli uniti disegni, nei quali:

15 -la figura 1 mostra un apparato in accordo con una forma di realizzazione della presente invenzione;

-le figure 2A, 2B, e 2C mostrano rappresentazioni grafiche dell’applicazione di rispettive strategie di riconoscimento, in accordo con la presente invenzione.

20 Nel corso della presente invenzione, si farà uso di alcuni termini, i quali vengono di seguito definiti.

Con “immagine digitale” si intende un’immagine che si presenti in formato digitale, indipendentemente dal modo con cui è stata acquisita per la prima volta; tale acquisizione sarà vantaggiosamente realizzata mediante una  
25 macchina fotografica digitale, ma altri metodi sono possibili, come indicato

nella parte introduttiva. Il soggetto dell'immagine sarà chiaro dal contesto.

Con "oggetto" si intende un qualunque elemento raffigurabile in un immagine digitale, ad esempio, bottiglie, lattine, barattoli, ma anche, libri ad 1 anta, libri a 2 ante, opuscoli, giornali, riviste, buste, insegne tavole sinottiche, ecc..

5           Con "testo presente nell'immagine" si intendono porzioni dell'immagine digitale contenenti rappresentazioni grafiche di elementi di testo, siano essi simboli alfanumerici o altri caratteri testuali di tutte le lingue incluse quelle asiatiche, arabe e cirilliche, ma anche simboli commercialmente noti, quali ad esempio  $\odot$ ,  $\text{er}$ ,  $\text{E}$ ,  $\text{e}$ ,  $\text{e}$ , lavare a mano, pura lana vergine, ecc.; in genere, si  
10 possono considerare "testo presente nell'immagine" tutto ciò che sia riconducibile ad uno dei caratteri leggibili da un computer come "testo" o comunque riconducibili a glifi.

Con "mezzi OCR" si intendono mezzi in grado di analizzare un'immagine digitale e riconoscere la presenza e la posizione di eventuale testo  
15 presente nell'immagine e di convertirlo in testo.

Tutto ciò premesso, con riferimento alle annesse figure, è rappresentato un apparato 1 per la determinazione del testo 2 contenuto in una immagine 3 raffigurante un oggetto 4.

Tale apparato 1 comprende:

20           -un dispositivo di elaborazione 5 avente mezzi di elaborazione 6 per elaborare l'immagine digitale 3 in modo da produrre una prima immagine modificata 3A;

          -mezzi 7 per eseguire un riconoscimento ottico di carattere (OCR) su detta prima immagine modificata 3A per determinare almeno una regione 3B  
25 contenente testo all'interno di detta immagine modificata 3A;

-mezzi 8 per convertire detto testo contenuto nella regione 3B in un primo segnale voce S1, detto segnale voce S1 essendo emettibile attraverso una interfaccia audio 9 operativamente connessa a detto dispositivo di elaborazione 1.

5 Tali mezzi per eseguire un riconoscimento ottico di carattere 7, si concretizzano in uno scanner oppure, vantaggiosamente, in una macchina fotografica (come rappresentato in figura 1), in una webcam, in uno scanner a planetario o in una videocamera digitale.

Giova rilevare che siffatti mezzi per eseguire un riconoscimento ottico di  
10 carattere 7 prevedono di associare alla prima regione 3B un primo valore di attendibilità C1, ossia un valore per mezzo del quale è possibile avere una stima della bontà del testo contenuto in tale prima regione 3B, nonché la posizione di tale regione 3B con riferimento alla prima immagine modificata 3A.

Giova rilevare che i mezzi 7 per eseguire un riconoscimento ottico di  
15 carattere (OCR) sono operativamente connessi con detto dispositivo di elaborazione 1, ad esempio, mediante un protocollo di comunicazione Bluetooth, mediante USB, mediante un protocollo di comunicazione Wi-Fi, via cavo, Wireless ecc..

In particolare, qualora si utilizzi una macchina fotografica digitale per  
20 acquisire l'immagine digitale 3 è bene sottolineare che oltre a trasferire le informazioni proprie relative a tale immagine digitale 3 è possibile che siano trasferite anche tutte quelle informazioni che la macchina fotografica memorizza in un cosiddetto "file exif" (o analoghi formati proprietari), ossia dettagli che contengono dettagli sulla macchina fotografica, sui parametri dello  
25 scatto quali, ad esempio, distanza di messa a fuoco, diaframma, tempo di

esposizione, focale, ecc. e sulla configurazione della masterizzazione.

In particolare, il dispositivo di elaborazione 1 si concretizza in un computer, sia esso un PC, un portatile, un palmare, un PDA o un dispositivo per telecomunicazioni come ad esempio un telefono cellulare.

5 Vantaggiosamente, in tale dispositivo di elaborazione 1 è installato un programma in grado di riconoscere il testo contenuto nella regione 3B dell'immagine modificata.

A tale fine i detti mezzi per elaborare 6 sono in grado di:

(a) elaborare detta immagine digitale 3 per produrre una prima  
10 immagine modificata 3A;

(b) eseguire un riconoscimento ottico di carattere (OCR) su detta prima immagine modificata 3A per determinare almeno una regione contenente testo 3B all'interno di detta immagine modificata 3A ed associare a detta prima regione un primo valore di attendibilità C1;

15 (c) convertire detto testo riconosciuto in un primo segnale voce S1; in cui detta fase di elaborare (a) comprende l'ulteriore fase di correggere in modo automatico i parametri grafici di detta immagine, detti parametri grafici comprendendo luminosità, contrasto saturazione e/o spazi colore equivalenti.

I parametri grafici possono essere modificati anche grazie alle  
20 informazioni contenute nei parametri "exif" di detta immagine 3A.

Giova rilevare che tale modifica automatica può anche avvenire modificando uno o più dei suddetti parametri luminosità, contrasto e/o saturazione o in spazi colore equivalenti.

Inoltre, tali mezzi per elaborare 6 sono anche in grado di espletare le  
25 seguenti fasi:

(d) confrontare detto primo valore di attendibilità C1 con un prefissato valore di attendibilità C2;

(e) selezionare almeno una strategia  $X_i$  di riconoscimento tra una pluralità "n" di strategie di riconoscimento X qualora detto primo valore di attendibilità C1 sia inferiore a detto prefissato valore di attendibilità C2.

Giova rilevare che ciascuna di detta pluralità di strategie X di riconoscimento è in grado di modificare i parametri grafici e geometrico/prospettici di detta prima immagine modificata 3A per generare una seconda immagine modificata 3C per incrementare detto primo valore C1.

10 Vantaggiosamente, le fasi precedentemente elencate (e), (b) e (c) possono essere ripetute fintanto che detto primo valore di attendibilità C1 è prossimo o superiore a detto prefissato valore di attendibilità C2.

È bene sottolineare che ciascuna  $X_i$  di detta pluralità di strategie X di riconoscimento è in grado di modificare i parametri grafici della prima  
15 immagine modificata 3A per generare la seconda 3C immagine modificata, iterando uno o più di detta pluralità di strategie X fino al raggiungimento di un valore di attendibilità C compatibile ai fine della presente invenzione.

In altre parole, a seguito all'applicazione sull'immagine modificata 3A di una o più di tali strategie X è possibile giungere ad un valore di affidabilità  
20 sufficiente elevato perché il testo contenuto nella regione di testo 3B della immagine modificata 3A possa essere letto con una precisione sufficiente così da generare il segnale voce S1 comprensibile all'utente.

Verranno ora illustrate, anche con riferimento alle figure da 2A a 2D, le suddette strategie X per ottenere il riconoscimento del testo contenuto nella  
25 regione 3B.

Una prima strategia  $X_1$  di riconoscimento comprendere la fase di:

(f) selezionare tra un elenco predefinito la tipologia cui appartiene detto oggetto 4, detto elenco comprendendo oggetti riconducibili a qualsiasi primitiva geometrica aventi forma cilindrica, sferica, conica, o anche oggetti riconducibili a qualsiasi primitiva prospettica singola, doppia e/o composta avente forma parallelepipedica.

Giova rilevare che la selezione può essere anche impostata manualmente da parte dell'utente specificando la specifica primitiva geometrica o primitiva prospettica cui appartiene l'oggetto 4.

10 Qualora l'oggetto 4 appartenga alla tipologia di oggetti aventi forma cilindrica, sferica e/o conica, la strategia  $X_1$  prevede che detta fase (f) comprenda l'ulteriore fase (f.1) o strategia  $X_{1,1}$  di applicare una funzione di trasformazione cilindrica inversa a detta regione di testo 3B.

15 Infatti, anche con riferimento alla figura 2A, grazie a tale funzione di trasformazione cilindrica inversa si ha che le linee di base del testo contenuta in tale regione 3B, che risultano essere curvate a causa della prospettiva, vengano ricondotte attraverso detta trasformazione inversa a linee rette 3D.

20 Qualora le aree di testo (ad esempio come rappresentato in figura 2d) sono due 3E e 3F, entrambe con bassa affidabilità, sui due lati dell'oggetto 4, e' probabile che si debba ruotare l'oggetto 4 di un predeterminato angolo, ad esempio, di 90 gradi.

A tale fine, la fase (f) comprende l'ulteriore fase (f.2) o strategia  $X_{1,2}$  di imporre una predefinita rotazione angolare attorno all'asse verticale di detto oggetto 4.

25 Vantaggiosamente è previsto che venga generato un secondo segnale

voce S2 per specificare la predefinita rotazione.

In particolare, i mezzi OCR 6 restituiscono come valore delle porzioni della immagine modificata 3A cui sono associati posizione e affidabilità "generale".

5           Ulteriormente all'interno di tali porzioni della immagine modificata 3° restituite dai mezzi OCR è anche associata a ciascun carattere la loro posizione e l'affidabilità.

Quindi nel caso in cui l'oggetto 4 raffigurato nell'immagine 3 e che abbia due colonne con affidabilità molto bassa, è probabile che tale oggetto 4  
10 appartenga ad una primitiva geometrica dell'elenco e che tale oggetto 4 sia stato fotografato dal lato sbagliato.

Il segnale voce S2 avverte l'utente e propone di effettuare la rotazione angolare attorno all'asse verticale dell'oggetto 4, ad esempio una rotazione pari a 90°.

15           Qualora invece l'oggetto 4 raffigurato nell'immagine 3 abbia due colonne con affidabilità medio alta, è probabile che tale oggetto 4 appartenga ad una primitiva geometrica dell'elenco e che tale oggetto 4 sia stato fotografato dal lato giusto. In tale scenario l'oggetto 4 risulta essere largo quanto la somma delle larghezze delle colonne.

20           È bene rilevare che la larghezza dell'oggetto 4 è calcolata come la larghezza della singola colonna o la somma delle due colonne.

Il metodo inventivo nel corso della strategia  $X_{1,1}$  e  $X_{1,2}$  imposta la funzione di trasformazione cilindrica inversa al centro della larghezza e dell'altezza dell'oggetto 4.

25           L'oggetto 4 è riconducibile a dette primitive geometriche è scelto tra il

gruppo comprendente bottiglie, lattine, barattoli e simili.

Qualora l'oggetto 4 appartenga alla tipologia di oggetti riconducibili a primitive prospettiche singola o doppia aventi forma parallelepipedica del tipo a libro ad una anta (figura 2B) o due ante (figura 2C), la fase (f) comprende l'ulteriore fase (f.3) o strategia  $X_{1,3}$  di applicare una funzione di prospettiva inversa a detta regione di testo.

Infatti, grazie a tale funzione di prospettiva inversa le linee di base del testo contenute nella regione 3b, curvate dalla struttura della relativa primitiva geometrica vengono ricondotte attraverso una trasformazione inversa a linee rette, come raffigurato nella zona 3D.

L'oggetto 4 è riconducibile a dette primitive prospettiche avente forma parallelepipedica è scelto tra il gruppo comprendete libri ad 1 anta, libri a 2 ante, opuscoli, giornali, riviste, buste, insegne tavole sinottiche.

Una ulteriore strategia  $X_{1,4}$ , eseguibile successivamente o indipendentemente dalle strategie prima descritte, comprende l'ulteriore fase (f.4) di interpolare la risoluzione nativa di detta prima immagine modificata 3A per generare detta seconda immagine modificata 3C avente una risoluzione superiore alla risoluzione nativa di detta prima immagine modificata.

Vantaggiosamente grazie a tale strategia  $X_{1,4}$  è possibile trovare caratteri con corpo piccolo internamente a detta regione di testo 3B.

Una ulteriore strategia  $X_{1,5}$ , eseguibile successivamente o indipendentemente dalle strategie prima descritte, comprende l'ulteriore fase (f.5) di applicare un primo filtro vettoriale del tipo di erosione del bordo detta regione contenente testo.

Vantaggiosamente grazie a tale strategia  $X_{1,5}$  è possibile distinguere le

scritte in grassetto internamente a detta regione di testo 3B.

Una ulteriore strategia  $X_{1,6}$ , eseguibile successivamente o indipendentemente dalle strategie prima descritte, comprende l'ulteriore fase (f.6) di applicare un secondo filtro vettoriale o grafico del tipo di contrasto massimo a detta regione contenente testo, e/o strategia  $X_{1,7}$  l'ulteriore fase (f.7) di applicare un terzo filtro vettoriale o grafico del tipo di contrasto minimo è a detta regione contenente testo.

Una ulteriore strategia  $X_{1,8}$ , eseguibile successivamente o indipendentemente dalle strategie prima descritte, comprende l'ulteriore fase (f.8) di applicare un quarto filtro vettoriale o grafico avente un primo valore di soglia per filtrare i pixel da bianchi a neri di detta regione contenente testo 3B, e/o strategia  $X_{1,9}$  di applicare un quinto filtro avente un secondo valore di soglia per filtrare i pixel da neri a bianchi di detta regione contenente testo.

È bene rilevare che detto valore di soglia delle strategie  $X_{1,8}$  e  $X_{1,9}$  è determinabile in funzione del valore di attendibilità prefissato.

Una seconda strategia di riconoscimento  $X_2$ , eseguibile successivamente o indipendentemente dalle strategie prima descritte, che comprende la fase di interpolare la risoluzione nativa di detta prima immagine modificata per generare detta seconda immagine modificata avente una risoluzione superiore alla risoluzione nativa di detta prima immagine modificata, come già descritto in precedenza.

Ulteriormente vi può essere una terza strategia di riconoscimento  $X_3$ , eseguibile successivamente o indipendentemente dalle strategie prima descritte, che comprende la fase di applicare detto primo filtro di erosione del bordo a detta regione contenente testo, come già descritto in precedenza.

Una quarta strategia di riconoscimento  $X_4$ , eseguibile successivamente o indipendentemente dalle strategie prima descritte, comprende la fase di applicare detto secondo filtro di contrasto massimo a detta regione contenente testo 3B.

- 5           Una quinta strategia di riconoscimento  $X_5$  comprende la fase di applicare detto terzo filtro di contrasto minimo a detta regione contenente testo 3B.

Una sesta strategia di riconoscimento  $X_6$  comprende la fase di applicare detto quarto filtro avente un primo valore di soglia per filtrare i pixel da bianchi  
10 a neri di detta regione contenente testo 3B.

Alternativamente, la sesta strategia di riconoscimento  $X_6$  comprende la fase di applicare detto quinto filtro avente un secondo valore di soglia per filtrare i pixel da neri a bianchi di detta regione contenente testo.

È bene rilevare che detto valore di soglia per filtrare i pixel da neri a  
15 bianchi e/o da bianchi a neri è determinabile in funzione del valore di attendibilità prefissato.

Giova altresì rilevare che detta almeno una regione contenente testo all'interno di detta immagine modificata qualora detto primo valore di attendibilità sia prossimo a zero coincide con detta immagine modificata.

20           Come si può apprezzare da quanto descritto, il metodo e l'apparato secondo l'invenzione permettono di soddisfare le esigenze di cui si è detto nella parte introduttiva della presente descrizione e di ovviare agli inconvenienti dei metodi e apparati della tecnica nota.

Ovviamente un tecnico del ramo, allo scopo di soddisfare esigenze  
25 contingenti e specifiche, potrà apportare numerose modifiche e varianti alle

configurazioni sopra descritte, tutte peraltro contenute nell'ambito di protezione dell'invenzione quale definita dalle seguenti rivendicazioni.

\*\*\* \* \*\*\*

**Ing. Aurelio PERANI**  
N. Iscr. ALBO 277  
(in proprio e per gli altri)

## RIVENDICAZIONI

1. Metodo per riconoscere testo in una immagine (3), in cui in detta immagine (3) è raffigurato un oggetto (4), **comprendente le fasi di:**

(a) elaborare una immagine digitale (3) per produrre una prima  
5 immagine modificata (3A);

(b) eseguire un riconoscimento ottico di carattere (OCR) su detta prima immagine modificata (3A) per determinare almeno una regione contenente testo (3B) all'interno di detta immagine modificata (3A) ed associare a detta prima regione (3B) un primo valore di attendibilità (C1);

10 (c) convertire detto testo riconosciuto in un primo segnale voce (S1);  
in cui detta fase di elaborare (a) comprende l'ulteriore fase di correggere in modo automatico i parametri grafici di detta immagine (3), detti parametri grafici comprendendo luminosità, contrasto saturazione e/o spazi colore equivalenti.

15 2. Metodo per riconoscere testo in una immagine in accordo con la rivendicazione 1, comprendente le ulteriori fasi di:

(d) confrontare detto primo valore di attendibilità (C1) con un prefissato valore di attendibilità (C2);

(e) selezionare almeno una strategia di riconoscimento ( $X_1; \dots, X_6$ ) tra  
20 una pluralità di strategie di riconoscimento qualora detto primo valore di attendibilità (C1) sia inferiore a detto prefissato valore di attendibilità (C2), ciascuna di detta pluralità di strategie di riconoscimento ( $X_1; \dots, X_6$ ) essendo atte a modificare i parametri grafici e geometrico/prospettici di detta prima immagine modificata (3A) per generare una seconda immagine modificata (3C)  
25 per incrementare detto primo valore di affidabilità (C1).

3. Metodo per riconoscere testo in una immagine in accordo con la rivendicazione 2, in cui detto metodo prevede di ripetere le fasi (e), (b) e (c) fintanto che detto primo valore di attendibilità (C1) è prossimo o superiore a detto prefissato valore di attendibilità (C2).
- 5 4. Metodo per riconoscere testo in una immagine in accordo con una qualunque delle rivendicazioni da 1 a 3, in cui una prima strategia ( $X_1$ ) di riconoscimento comprendere la fase di:
- (f) selezionare tra un elenco predefinito la tipologia cui appartiene detto oggetto (4), detto elenco comprendendo oggetti riconducibili a qualsiasi primitiva geometrica aventi forma cilindrica, sferica, conica, o anche oggetti riconducibili a qualsiasi primitiva prospettica singola, doppia e/o composta avente forma parallelepipedica.
- 10
5. Metodo per riconoscere testo in una immagine in accordo con la rivendicazione 4, in cui qualora detto oggetto (4) appartenga alla tipologia di oggetti aventi forma cilindrica, sferica e/o conica detta fase (f) comprende ( $X_{1,1}$ ) l'ulteriore fase (f.1) di applicare una funzione di trasformazione cilindrica inversa a detta regione di testo.
- 15
6. Metodo per riconoscere testo in una immagine in accordo con la rivendicazione 4 o 5, in cui detta fase (f) comprende l'ulteriore fase (f.2) di imporre ( $X_{1,2}$ ) una predefinita rotazione angolare attorno all'asse verticale di detto oggetto.
- 20
7. Metodo per riconoscere testo in una immagine in accordo con la rivendicazione 6, in cui detta fase (f.2) prevede di generare un secondo segnale voce (S2) per specificare detta predefinita rotazione angolare.
- 25 8. Metodo per riconoscere testo in una immagine in accordo con la

- rivendicazione 4, in cui qualora detto oggetto (4) appartenga alla tipologia di oggetti riconducibili a primitive prospettiche singola o doppia aventi forma parallelepipeda, detta fase (f) comprende l'ulteriore fase (f.3) di applicare ( $X_{1.3}$ ) una funzione di prospettiva inversa a detta regione di testo (3B).
- 5    **9.** Metodo per riconoscere testo in una immagine in accordo con una qualunque delle rivendicazioni da 4 a 8, in cui detta fase (f) comprende l'ulteriore fase (f.4) di interpolare ( $X_{1.4}$ ) la risoluzione nativa di detta prima immagine modificata per generare detta seconda immagine modificata avente una risoluzione superiore alla risoluzione nativa di detta prima immagine modificata (3B).
- 10   **10.** Metodo per riconoscere testo in una immagine in accordo con una qualunque delle rivendicazioni da 4 a 9, in cui detta fase (f) comprende l'ulteriore fase (f.5) di applicare ( $X_{1.5}$ ) un primo filtro vettoriale o grafico del tipo di erosione del bordo detta regione contenente testo (3B).
- 15   **11.** Metodo per riconoscere testo in una immagine in accordo con una qualunque delle rivendicazioni da 4 a 10, in cui detta fase (f) comprende l'ulteriore fase (f.6) di applicare ( $X_{1.6}$ ) un secondo filtro vettoriale o grafico del tipo di contrasto massimo a detta regione contenente testo (3B).
- 20   **12.** Metodo per riconoscere testo in una immagine in accordo con una qualunque delle rivendicazioni da 4 a 10, in cui detta fase (f) comprende l'ulteriore fase (f.7) di applicare ( $X_{1.7}$ ) un terzo filtro vettoriale o grafico del tipo di contrasto minimo a detta regione contenente testo (3B).
- 25   **13.** Metodo per riconoscere testo in una immagine in accordo con una qualunque delle rivendicazioni da 4 a 12, in cui detta fase (f) comprende l'ulteriore fase (f.8) di applicare ( $X_{1.8}$ ) un quarto filtro vettoriale o grafico avente un primo valore di soglia per filtrare i pixel da bianchi a neri di detta

regione contenente testo (3B).

14. Metodo per riconoscere testo in una immagine in accordo con una qualunque delle rivendicazioni da 4 a 12, in cui detta fase (f) comprende l'ulteriore fase (f.9) di applicare ( $X_{1,9}$ ) un quinto filtro vettoriale o grafico  
5  
avente un secondo valore di soglia per filtrare i pixel da neri a bianchi di detta regione contenente testo (3B).

15. Metodo per riconoscere testo in una immagine in accordo con le rivendicazioni 13 e 14, in cui detto valore di soglia è determinabile in funzione del valore di attendibilità prefissato (C2).

10  
16. Metodo per riconoscere testo in una immagine in accordo con una qualunque delle rivendicazioni da 1 a 3, in cui una seconda strategia di riconoscimento ( $X_2$ ) comprendere la fase di interpolare la risoluzione nativa di detta prima immagine modificata (3A) per generare detta seconda immagine modificata (3C) avente una risoluzione superiore alla risoluzione nativa di detta  
15  
prima immagine modificata (3A).

17. Metodo per riconoscere testo in una immagine in accordo con una qualunque delle rivendicazioni da 1 a 4 o 16, in cui una terza strategia di riconoscimento ( $X_3$ ) comprendere la fase di applicare detto primo filtro di erosione del bordo a detta regione contenente testo (3B).

20  
18. Metodo per riconoscere testo in una immagine in accordo con una qualunque delle rivendicazioni da 1 a 4 o da 16 a 17, in cui una quarta strategia di riconoscimento ( $X_4$ ) comprende la fase di applicare detto secondo filtro di contrasto massimo a detta regione contenente testo (3B).

19. Metodo per riconoscere testo in una immagine in accordo con una  
25  
qualunque delle rivendicazioni da 1 a 4 o da 16 a 18, in cui una quinta strategia

di riconoscimento ( $X_5$ ) comprende la fase di applicare detto terzo filtro di contrasto minimo è a detta regione contenente testo (3B).

20. Metodo per riconoscere testo in una immagine in accordo con una qualunque delle rivendicazioni da 1 a 4 o da 16 a 19, in cui una sesta strategia di  
5 riconoscimento ( $X_6$ ) comprende la fase di applicare detto quarto filtro avente un primo valore di soglia per filtrare i pixel da bianchi a neri di detta regione contenente testo (3B).

21. Metodo per riconoscere testo in una immagine in accordo con una qualunque delle rivendicazioni da 1 a 4 o da 16 a 19, in cui una sesta strategia di  
10 riconoscimento ( $X_6$ ) comprende la fase di applicare detto quinto filtro avente un secondo valore di soglia per filtrare i pixel da neri a bianchi di detta regione contenente testo (3B).

22. Metodo per riconoscere testo in una immagine in accordo con le rivendicazioni 20 e 21, in cui detto valore di soglia è determinabile in funzione  
15 del valore di attendibilità prefissato.

23. Metodo per riconoscere testo in una immagine in accordo con la rivendicazione 4, in cui detto oggetto riconducibile a dette primitive geometriche è scelto tra il gruppo comprendente bottiglie, lattine, barattoli.

24. Metodo per riconoscere testo in una immagine in accordo con la  
20 rivendicazione 8, in cui detto oggetto a dette primitive prospettiche avente forma parallelepipedica è scelto tra il gruppo comprendete libri ad 1 anta, libri a 2 ante, opuscoli, giornali, riviste, buste, insegne tavole sinottiche.

25. Metodo per riconoscere testo in una immagine in accordo con una qualunque delle precedenti rivendicazioni da 1 a 24, in cui detta almeno una  
25 regione contenente testo (3B) all'interno di detta immagine modificata (3A)

qualora detto primo valore di attendibilità (C1) sia prossimo a zero coincide con detta immagine modificata (3A).

26. Prodotto informatico direttamente caricabile nella memoria di un computer, comprendente porzioni di codice di programma suscettibile di realizzare il  
5 metodo in accordo con una qualunque delle rivendicazioni da 1 a 25 quando fatto girare su detto computer.

27. Apparato per riconoscere testo in una immagine (3), in cui in detta immagine (3) è raffigurato un oggetto (4), detto apparato comprendente:

-un dispositivo di elaborazione (1) avente mezzi per elaborare una immagine  
10 digitale (3) per produrre una prima immagine modificata (3A);

-mezzi (7) per eseguire un riconoscimento ottico di carattere (OCR) su detta prima immagine modificata (3A) per determinare almeno una regione contenente testo (3B) all'interno di detta immagine modificata (3A) e per associare a detta prima regione (3B) un primo valore di attendibilità (C1);

15 -mezzi (8) per convertire detto testo riconosciuto in un primo segnale voce (S1), detto segnale voce (S1) essendo emettibile attraverso una interfaccia audio (9) operativamente connessa a detto dispositivo di elaborazione (1);

**caratterizzato dal fatto** che detti mezzi (7) per eseguire un riconoscimento ottico di carattere comprendono un programma in accordo con una qualunque  
20 delle rivendicazioni da 1 a 25.

28. Apparato per riconoscere testo in una immagine, in accordo con la rivendicazione 27, in cui detti mezzi (7) per acquisire detta immagine digitale sono operativamente connessi con detto dispositivo di elaborazione (1).

29. Apparato per riconoscere testo in una immagine, in accordo con una  
25 qualunque delle rivendicazioni da 27 a 28, in cui detti mezzi (7) per eseguire un

riconoscimento ottico di carattere (OCR) su detta prima immagine modificata (3A) sono atti a generare un primo valore di attendibilità (C1) per detto almeno una regione contenente testo (3B).

- 5 30. Apparato per riconoscere testo in una immagine, in accordo con la rivendicazione 27, in cui in cui detti mezzi per elaborare sono atti a confrontare detto primo valore di attendibilità (C1) con un prefissato valore di attendibilità (C2) e a selezionare almeno una strategia di riconoscimento ( $X_1; \dots, X_6$ ) tra una pluralità di strategie di riconoscimento qualora detto primo valore di attendibilità (C1) sia inferiore a detto prefissato valore di attendibilità (C2),
- 10 ciascuna di detta pluralità di strategie di riconoscimento ( $X_1; \dots, X_6$ ) essendo atte a modificare i parametri grafici di detta prima immagine modificata (3A) per generare una seconda immagine modificata (3C) fintanto che detto primo valore di attendibilità (C1) è prossimo o superiore a detto prefissato valore di attendibilità (C2).
- 15 31. Apparato per riconoscere testo in una immagine, in accordo con una qualunque delle rivendicazioni da 27 a 30, in cui detto dispositivo di elaborazione (1) comprende un computer, un notebook, un palmare, un PDA o un sistema di telecomunicazioni quali un telefono cellulare.

**Ing. Aurelio PERANI**  
N. Iscr. ALBO 277  
(in proprio e per gli altri)

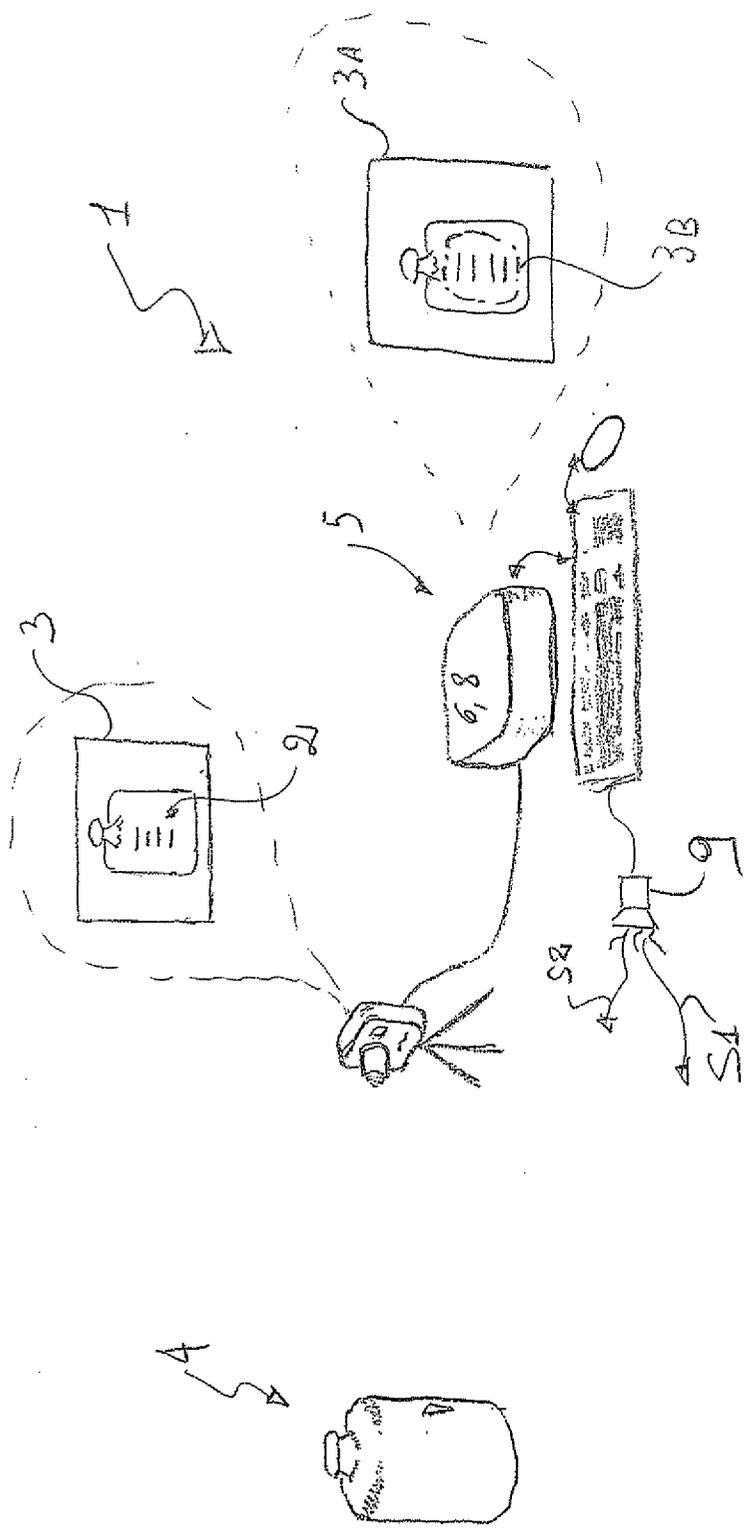


Fig. 1

