

ALMA MATER STUDIORUM - UNIVERSITÀ DI BOLOGNA

FACOLTA' DI INGEGNERIA

CORSO DI LAUREA IN INGEGNERIA ELETTRONICA

TESI DI LAUREA

in
Geometria e Algebra

Analisi comparata di motori di ricerca per immagini

CANDIDATO:
Giovanni Di Stanislao

RELATORE:
Massimo Ferri

Anno Accademico 2013/2014
Sessione II

Ai miei genitori.

Indice

Introduzione.....	4
1 Metodologie.....	5
1.1 Colore.....	5
1.2 Texture.....	6
1.2.1 Caratteristiche di Tamura.....	7
1.2.1.1 Grossolanità.....	7
1.2.1.2 Contrasto.....	8
1.2.1.3 Direzionalità.....	8
1.2.2 Caratteristiche di Wold.....	9
1.2.3 Modello simultaneo auto-regressivo.....	9
1.2.4 Filtri di Gabor.....	9
1.2.5 Trasformata wavelet.....	11
1.3 Forma.....	12
1.3.1 Momenti invarianti.....	12
1.3.2 Angoli di rotazione.....	13
1.3.3 Descrittori di Fourier.....	14
1.3.4 Compattezza.....	15
1.3.5 Asse maggiore.....	15
1.3.6 Asse minore.....	15
1.3.7 Rettangolo di base.....	16

1.3.8	Eccentricità.....	16
1.3.9	Allungamento.....	16
1.4	Informazioni spaziali.....	16
2	Motori di ricerca	
2.1	Google.....	21
2.1.1	Interfaccia utente.....	21
2.1.2	Immagini visivamente simili.....	22
2.1.3	Risultati di Google per le immagini test.....	35
2.2	TinEye.....	48
2.2.1	Interfaccia utente.....	48
2.2.2	Risultati di TinEye per le immagini test.....	49
2.3	Yandex.....	52
2.3.1	Interfaccia utente.....	52
2.3.2	Risultati di Yandex per le immagini test.....	53
	Conclusione.....	59
	Bibliografia.....	60

Introduzione

La diffusione esponenziale delle grandi collezioni di immagini digitali ha dato luogo ad un interesse crescente per i tool automatici in grado di memorizzare, organizzare e recuperare in modo efficiente le immagini. Questo ha reso possibile la crescita di una promettente area di ricerca conosciuta come Content-Based Image Retrieval (CBIR) che coinvolge la progettazione e lo sviluppo di sistemi capaci di utilizzare delle caratteristiche visuali come texture, colore e forma per il processo di indicizzazione e recupero delle immagini.

In questa tesi si osserveranno tre motori di ricerca per immagini: Google, TinEye e Yandex e si procederà a una descrizione qualitativa del loro funzionamento confrontando quali dei tre sia il più performante. Sono stati scelti principalmente per la grande quantità di immagini nel database.

Nel capitolo uno saranno introdotte le metodologie che un motore di ricerca può utilizzare per il recupero delle immagini. Nel capitolo due si testeranno i motori di ricerca.

Capitolo 1

Metodologie

Un'immagine può possedere proprietà di carattere sia semantico che visivo. Il contenuto semantico è ottenuto o da annotazioni testuali o da complesse procedure deduttive basate sul contenuto visivo. Il contenuto visivo può essere generale e include: colore, texture, forma, relazioni spaziali etc. o specifico del campo (es: volti umani). Ogni descrittore può avere carattere o locale o globale; globale se è riferito all'intera immagine e locale se è riferito a parti della stessa.

1.1 Colore

Per le immagini con sole tonalità di grigio, il colore è un vettore unidimensionale, mentre per le immagini a colori, è un vettore a tre dimensioni quindi, ogni pixel dell'immagine può essere rappresentato come un punto 3D nello spazio dei colori. Gli spazi di colori più comunemente usati sono RGB, CIE L^*a^*b , CIE L^*u^*v , HSV.

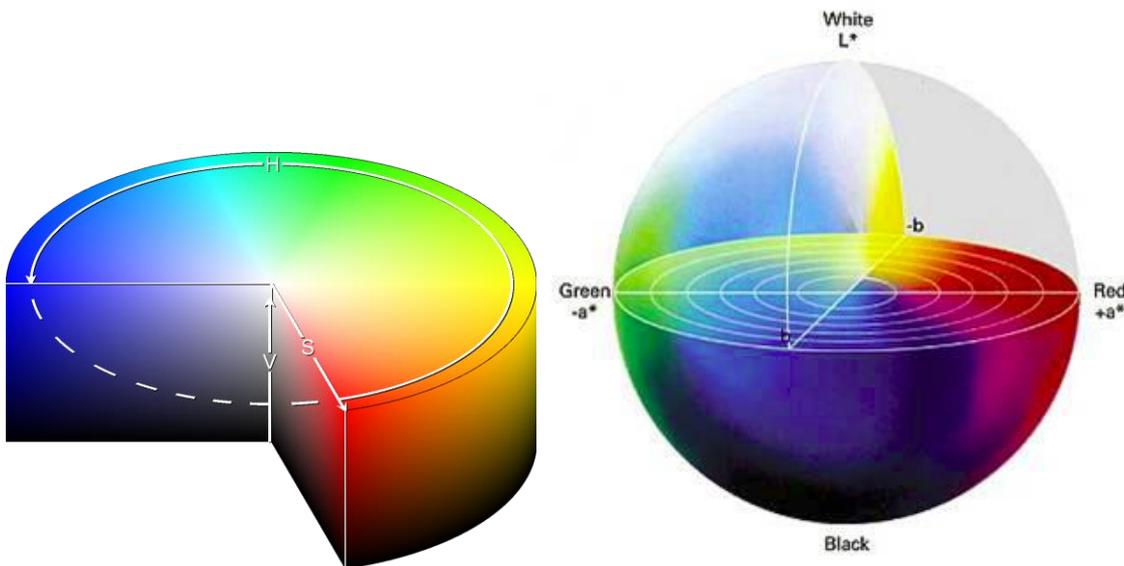


Fig.1 esempi di spazi di colore.

Ci sono diversi descrittori che caratterizzano il colore come i momenti, l'istogramma, il vettore di coerenza e il correlogramma. I momenti più usati sono quelli del primo, secondo e terzo ordine che rappresentano rispettivamente la media, la varianza e l'asimmetria [1].

$$\mu_i = \frac{1}{N} \sum_{j=1}^N f_{ij} \quad (\text{momento ordine I})$$

$$\sigma_i = \left(\frac{1}{N} \sum_{j=1}^N (f_{ij} - \mu_i)^2 \right)^{\frac{1}{2}} \quad (\text{momento ordine II})$$

$$s_i = \left(\frac{1}{N} \sum_{j=1}^N (f_{ij} - \mu_i)^3 \right)^{\frac{1}{3}} \quad (\text{momento ordine III})$$

Sono calcolati per ogni componente del vettore colore. L'istogramma è una rappresentazione della distribuzione dei colori in un'immagine. Il range dei colori che un pixel dell'immagine può assumere viene suddiviso in intervalli. Ad ogni intervallo viene associato il numero di pixel che appartengono a quell'intervallo. Il vettore di coerenza è un vettore che ha dimensione uguale al numero degli intervalli usati nell'istogramma e ad ogni suo elemento associa una coppia di elementi che corrispondono al numero di pixel coerenti o incoerenti. Un pixel viene definito coerente se appartiene a una larga regione colorata uniformemente, incoerente altrimenti. Il correlogramma indica la probabilità di trovare un pixel di colore j a una distanza k da un pixel di colore i nell'immagine.

1.2 Texture

Non c'è una definizione formale di texture, intuitivamente la texture descrive proprietà quali la grossolanità, la levigatezza e la regolarità di un'intera immagine o di regioni della stessa. Classificare una texture è più difficile che discriminarla poiché una stessa texture può apparire in un'immagine in modi anche assai diversi tra loro. Dipende da: grado di luminosità, orientazione, scala etc. Gli approcci principali per la descrizione della texture sono l'approccio statistico e l'approccio strutturale. Le tecniche strutturali descrivono una texture identificando texture primitive e le regole del loro posizionamento; sono usate principalmente quando le texture sono molto regolari.



Fig.2 Es di texture regolare.

I metodi statistici portano alla caratterizzazione delle strutture come lisce, ruvide, granulose, etc. essi sono:

1.2.1 Caratteristiche di Tamura

Le caratteristiche di Tamura, sono ideate in accordo con gli studi psicologici sulla percezione umana della texture; includono la grossolanità, il contrasto e la direzionalità [2].

1.2.1.1 Grossolanità (granularità)

Per calcolare la granularità della texture occorre calcolare le medie mobili $A_k(x, y)$ usando finestre di dimensioni $2^k \times 2^k$ ($k=0, 1, \dots, 5$) per ogni pixel (x, y) .

$$A_k(x, y) = \frac{1}{2^{2k}} \sum_{i=x-2^{k-1}}^{x+2^{k-1}-1} \sum_{j=y-2^{k-1}}^{y+2^{k-1}-1} g(i, j)$$

Dove $g(i, j)$ è l'intensità del pixel in (i, j) ;

Sono poi calcolate le differenze tra coppie di medie mobili non sovrapposte nelle direzioni orizzontali e verticali.

$$E_{k,h}(x, y) = |A_k(x + 2^{k-1}, y) - A_k(x - 2^{k-1}, y)|$$

$$E_{k,v}(x, y) = |A_k(x, y + 2^{k-1}) - A_k(x, y - 2^{k-1})|$$

Fatto ciò, si calcola il valore di k che massimizza E in entrambe le direzioni per avere la migliore dimensione della finestra per ogni pixel.

$$S_{best}(x, y) = 2^k$$

La grossolanità è ottenuta calcolando la media di S_{best} sull'intera immagine.

1.2.1.2 Contrasto

La formula per il calcolo del contrasto è la seguente:

$$F_{con} = \frac{\sigma}{\alpha_4^{\frac{1}{4}}}$$

Dove $\alpha_4 = \mu_4 / \sigma^4$. μ_4 è il momento quarto riferito alla media e σ^2 è la varianza.

1.2.1.3 Direzionalità



Fig.3 es di Texture con alta direzionalità a sinistra e casuale a destra.

Per calcolare la direzionalità, l'immagine è convoluta con 2 matrici 3x3

$$\begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{pmatrix}$$

e il vettore gradiente è calcolato per ogni pixel. Il suo modulo e la sua direzione sono così definiti

$$\Delta G = (|\Delta_H| + |\Delta_V|) / 2$$

$$\theta = \tan^{-1}(\Delta_V / \Delta_H) + \pi / 2$$

In cui Δ_H e Δ_V sono le differenze orizzontali e verticali di convoluzione.

Poi si quantizza θ e si contano i pixel il cui modulo del gradiente ad essi associato supera una certa soglia. Si costruisce quindi l'istogramma associato a θ .

Questo istogramma mostrerà forti picchi per immagini ad alta direzionalità e sarà relativamente piatto per immagini a bassa direzionalità.

1.2.2 Le caratteristiche di Wold

Le caratteristiche di Wold forniscono un altro approccio nella descrizione delle texture in termini di proprietà percettive [3]. Le tre componenti di Wold, armonica, evanescente e indeterministica corrispondono alla periodicità, direzionalità e casualità della texture.

1.2.3 Modello simultaneo auto-regressivo(SAR)

Nel modello SAR le intensità dei pixel sono considerate come variabili random [4]. L'intensità $g(x,y)$ del pixel (x,y) può essere considerata come combinazione lineare dei valori $g(x',y')$ dei pixel vicini a cui va aggiunto un termine additivo dovuto al rumore $\varepsilon(x,y)$.

$$g(x, y) = \mu + \sum_{(x',y') \in D} \theta(x', y')g(x', y') + \varepsilon(x, y).$$

In cui μ è determinato dalla media dell'intera immagine, D è l'insieme dei pixel vicini e $\theta(x',y')$ è il peso associato a ogni pixel vicino, $\varepsilon(x,y)$ è una variabile gaussiana casuale indipendente a media nulla e varianza σ^2 .

I parametri θ e σ sono usati per misurare la texture; un alto valore di σ implica una granularità fine, e alti valori ad esempio per $\theta(x, y + 1)$ e $\theta(x, y - 1)$ indicano che la texture è orientata verticalmente.

1.2.4 Filtri di Gabor

Osservazioni generali

Per analizzare segnali non-stazionari bisogna utilizzare una rappresentazione che sia locale sia nel dominio del tempo (o nello spazio), sia in quello delle frequenze. Per località nel tempo si intende la capacità di localizzare il contenuto di frequenze appartenenti ad un certo intervallo temporale. La trasformata di Fourier non permette la localizzazione temporale, poiché per calcolare il suo valore ad una determinata frequenza si deve conoscere l'intero segnale. Se in un certo istante il segnale viene alterato, questo si ripercuote sull'intero spettro. Un approccio per ottenere un'analisi tempo-frequenza locale consiste nel tagliare il segnale in parti e applicare un'analisi di Fourier su ciascuna di queste parti. Le funzioni, però, ottenute da una segmentazione di questo tipo, non smussata, non sono periodiche, poiché la trasformata di Fourier interpreta i salti agli estremi come discontinuità o come

brusche variazioni del segnale. Per evitare tali artefatti, è stato introdotto il concetto di Windowing (finestratura). Invece di localizzare il segnale mediante una funzione rettangolare, si usa per la segmentazione una funzione smussata, che assuma un valore vicino ad 1 intorno all'origine e decada verso il valore 0 agli estremi. Tale funzione è svolta dal filtro di Gabor [5].

Filtri di Gabor

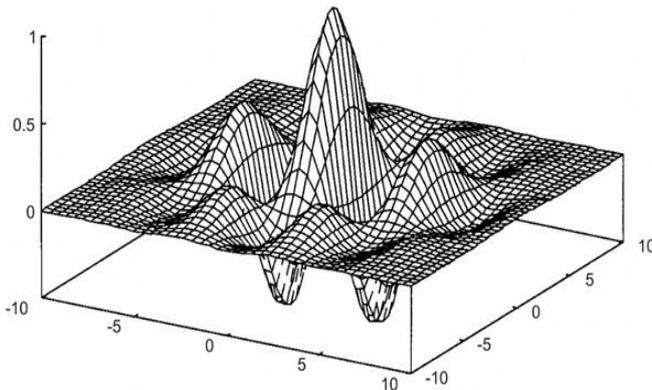


Fig.4 Filtro di Gabor

Nel dominio spaziale la funzione di Gabor $g(x,y)$ è definita come una funzione gaussiana modulata da un esponenziale orientato.

$$g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right) + 2\pi jWx\right]$$

In cui σ_x e σ_y sono le deviazioni standard della funzione di Gabor lungo le direzioni x e y.

Una serie di filtri di Gabor può essere ottenuta con opportune dilatazioni e rotazioni di $g(x,y)$ in modo da osservare l'orientazione della texture.

$$g_{mn}(x, y) = a^{-m} g(x', y')$$

$$x' = a^{-m}(x \cos\theta + y \sin\theta)$$

$$y' = a^{-m}(-x \sin\theta + y \cos\theta)$$

$a > 1$, $\theta = n\pi/K$, $n=0,1,\dots,K-1$, $m=0,1,\dots,S-1$. K e S sono il numero di orientazioni e scale.

1.2.5 La trasformata Wavelet

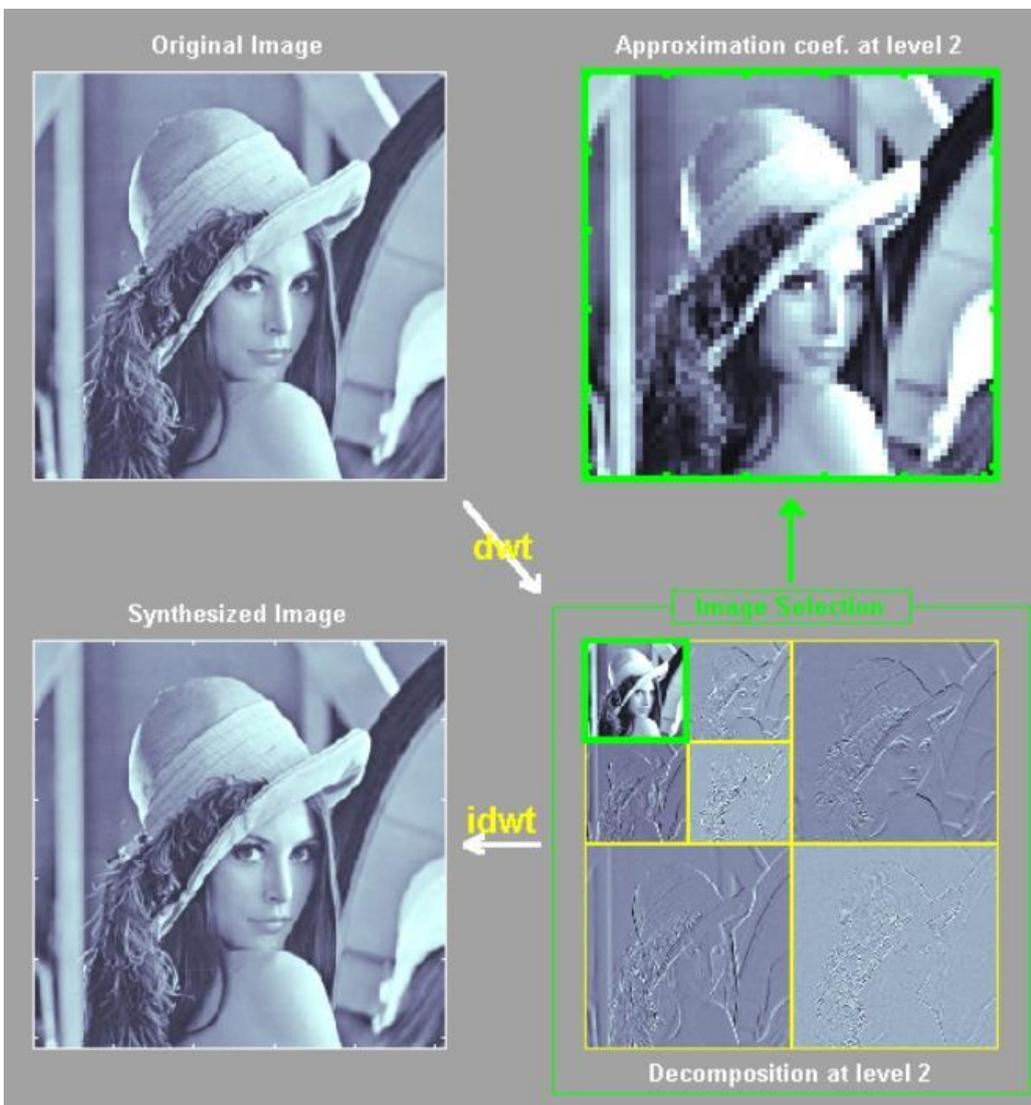
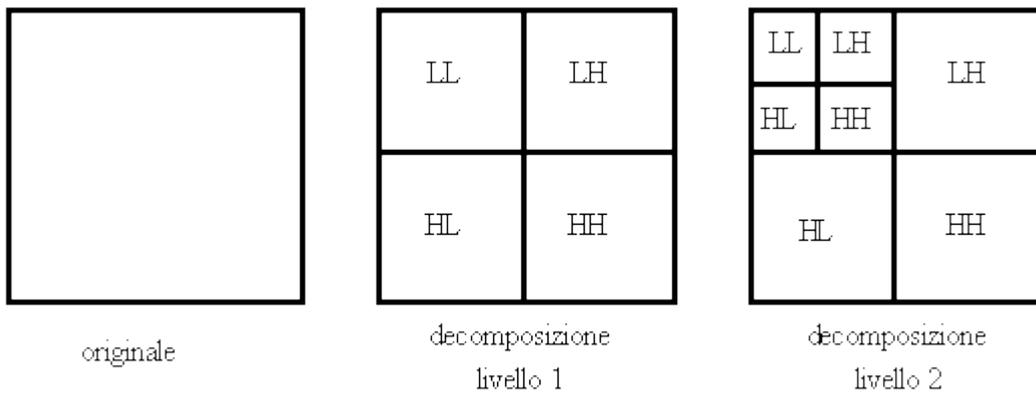


Fig.5 applicazione trasformata wavelet

La trasformata wavelet decompone il segnale attraverso una famiglia di funzioni base $\psi_{mn}(x)$ a media nulla ed energia finita, ottenute attraverso la traslazione e la dilatazione di una wavelet (piccola onda) madre [6]. Un esempio

$$\psi_{mn}(x) = 2^{-\frac{m}{2}} \psi(2^{-m}x - n)$$

In cui m ed n sono rispettivamente parametri di dilatazione e di traslazione. Quindi un segnale $f(x)$ può essere rappresentato come:

$$f(x) = \sum_{m,n} c_{mn} \psi_{mn}(x)$$

Il calcolo della trasformata wavelet di un segnale 2d coinvolge filtraggi ricorsivi e sotto-campionamenti. A ogni livello, il segnale è decomposto in 4 sotto-bande di frequenza: LL, LH, HL, HH dove L sta per bassa frequenza e H sta per alta frequenza.

Dopo la decomposizione, il features vector può essere costruito usando la media e la deviazione standard della distribuzione energetica di ogni sotto-banda per ogni livello.

1.3 Forma

Le caratteristiche di forma di oggetti o regioni sono usate in molti sistemi di recupero delle immagini basate sui contenuti [7]. Spesso vengono utilizzate dopo che l'immagine è stata divisa in regioni, ma poiché è difficile raggiungere una segmentazione dell'immagine accurata e robusta tali caratteristiche sono usate soltanto in applicazioni specifiche. Le proprietà salienti che deve avere un buon descrittore della forma di un oggetto sono l'invarianza alla rotazione, allo scaling e alla traslazione dell'oggetto.

1.3.1 Momenti invarianti

Per caratterizzare la forma di un oggetto viene usualmente utilizzato un insieme ristretto di momenti possibilmente invarianti rispetto alla posizione, orientazione e scala dell'oggetto. Se l'oggetto R è rappresentato da un'immagine binaria il momento centrale di ordine $p+q$ per la forma dell'oggetto R sono definiti da:

$$\mu_{p,q} = \sum_{(x,y) \in R} (x - x_c)^p (y - y_c)^q$$

In cui (x_c, y_c) è il baricentro dell'oggetto.

Questo momento centrale può essere normalizzato per essere indipendente dallo scaling:

$$\eta_{p,q} = \frac{\mu_{p,q}}{\mu_{0,0}^\gamma}, \gamma = \frac{p+q+2}{2}$$

Da questi momenti possiamo derivare momenti invarianti per traslazione, rotazione e cambiamento di scala:

$$\Phi_1 = \mu_{2,0} + \mu_{0,2}$$

$$\Phi_2 = (\mu_{2,0} + \mu_{0,2})^2 + 4\mu_{1,1}^2$$

$$\Phi_3 = (\mu_{3,0} - 3\mu_{1,2})^2 + (\mu_{0,3} - 3\mu_{2,1})^2$$

$$\Phi_4 = (\mu_{3,0} + \mu_{1,2})^2 + (\mu_{0,3} + \mu_{2,1})^2$$

$$\Phi_5 = (\mu_{3,0} - 3\mu_{1,2})(\mu_{3,0} + \mu_{1,2})[(\mu_{3,0} + \mu_{1,2})^2 - 3(\mu_{0,3} + \mu_{2,1})^2] +$$

$$+(\mu_{0,3} - 3\mu_{2,1})(\mu_{0,3} + \mu_{2,1})[(\mu_{0,3} + \mu_{2,1})^2 - 3(\mu_{3,0} + \mu_{1,2})^2]$$

$$\Phi_6 = (\mu_{2,0} + \mu_{0,2})[(\mu_{3,0} + \mu_{1,2})^2 - (\mu_{0,3} + \mu_{2,1})^2] + 4\mu_{1,1}(\mu_{3,0} + \mu_{1,2})(\mu_{0,3} + \mu_{2,1})$$

$$\Phi_7 = (3\mu_{2,1} - \mu_{0,3})(\mu_{3,0} + \mu_{1,2})[(\mu_{3,0} + \mu_{1,2})^2 - 3(\mu_{0,3} + \mu_{2,1})^2]$$

.

1.3.2 Angoli di rotazione

Il contorno di un oggetto 2D può essere rappresentato come una sequenza chiusa di pixel (x_s, y_s) di contorno, in cui $0 \leq s \leq N-1$ e N è il numero totale di pixel sul bordo dell'oggetto. La funzione $\Theta(s)$, che misura l'angolo della tangente in funzione della lunghezza d'arco s in relazione a un punto di riferimento sul bordo dell'oggetto può essere definita come:

$$\Theta(s) = \tan^{-1}\left(\frac{y'_s}{x'_s}\right)$$

$$y'_s = \frac{dy_s}{ds}$$

$$x'_s = \frac{dx_s}{ds}$$

Uno dei maggiori problemi con questa rappresentazione è che dipende dalla rotazione dell'oggetto e dalla scelta del punto di riferimento. Se noi scorriamo il punto di riferimento lungo il bordo di una quantità t , poi la nuova funzione diventerà $\Theta(s+t)$.

Se invece rotiamo l'oggetto di un angolo ω la nuova funzione sarà $\Theta(s)+\omega$.

Quindi per confrontare la somiglianza tra due oggetti A e B con tale funzione, si deve calcolare la distanza minima su tutti i possibili shift t e le rotazioni ω .

Es:

$$d_p(A, B) = \left(\min_{\omega \in \mathbb{R}, t \in [0, 1]} \int_0^1 | \theta_A(s + t) - \theta_B(s) + \omega |^p ds \right)^{\frac{1}{p}}$$

Si è supposto che ogni oggetto sia stato ridimensionato in modo che la lunghezza del suo perimetro sia uno.

1.3.3 Descrittori di Fourier

I descrittori di Fourier descrivono la forma di un oggetto con la trasformata di Fourier del suo contorno. Consideriamo ancora il contorno di un oggetto 2D come una sequenza chiusa di pixel di bordo (x_s, y_s) , in cui $0 \leq s \leq N-1$ e N è il numero totale di pixel sul bordo dell'oggetto. Poi tre tipi di rappresentazioni di bordo possono essere definite:

curvatura, distanza dal baricentro, e la funzione coordinata complesse.

La curvatura $K(s)$ a un punto s lungo il contorno è definita come un indice di cambiamento della direzione della tangente al contorno.

$$K(s) = \frac{d}{ds} \Theta(s)$$

Dove $\Theta(s)$ è la funzione definita precedentemente.

La distanza dal baricentro è definita come la distanza tra i pixel di bordo e il baricentro (x_c, y_c) dell'oggetto. $(y_s - y_c)$

$$R(s) = \sqrt{(x_s - x_c)^2 + (y_s - y_c)^2}$$

La coordinata complessa è ottenuta semplicemente rappresentando le coordinate del bordo come un numero complesso:

$$Z(s) = (x_s - x_c) + j(y_s - y_c)$$

La trasformata di Fourier di questi tre tipi di descrittori del contorno generano tre insiemi di coefficienti complessi che rappresentano la forma di un oggetto nel dominio delle frequenze. I coefficienti alle basse frequenze descrivono le proprietà generali della forma

dell'oggetto, mentre i coefficienti ad alte frequenze ne descrivono i dettagli. Per ottenere l'invarianza nella rotazione, sono usate solo l'ampiezza dei coefficienti complessi e le fasi sono scartate. Per ottenere una rappresentazione invariante allo scaling, l'ampiezza dei coefficienti sono divise per l'ampiezza della componente a frequenza zero o per l'ampiezza del primo coefficiente diverso da zero. L'invarianza per traslazione è ottenuta direttamente dalla rappresentazione del contorno.

1.3.4 Compattezza

Le misure geometriche perimetro ed area sono entrambi poco significative quando un oggetto è osservato da distanze diverse. In questo caso è necessario definire misure di forme invarianti al cambiamento di scala dell'oggetto nel piano immagine.

La compattezza C è una semplice misura (non dimensionale) di forma come :

$$C = \frac{P^2}{A}$$

Dove P ed A sono rispettivamente il perimetro e l'Area della regione.

Il cerchio è la figura geometrica più compatta nello spazio Euclideo con valore minimo della compattezza uguale a 4π .

1.3.5 Asse Maggiore

Si definisce asse maggiore di un oggetto la corda di massima lunghezza tracciata tra due punti estremi del contorno dell'oggetto stesso.

Se indichiamo con P_i e P_j due pixel del contorno C e con $d(P_i, P_j)$ la distanza tra i due punti, la lunghezza dell'asse maggiore D_1 è dato da:

$$D_1 = \max[d(P_i, P_j)] = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}$$

L'angolo θ formato tra l'asse maggiore e l'asse delle x definisce l'orientazione dell'oggetto nel piano immagine ed è calcolato da:

$$\theta = \arctan[(y_j - y_i)/(x_j - x_i)]$$

La lunghezza dell'asse maggiore D_1 e l'angolo di orientazione θ dell'asse maggiore rispetto all'asse x costituiscono due parametri utili di descrizione del contorno.

1.3.6 Asse minore

L'asse minore di un oggetto è definito come la corda di massima lunghezza che può essere tracciata in direzione perpendicolare all'asse maggiore.

La lunghezza dell'asse minore è calcolata analogamente a quella dell'asse maggiore considerando in questo caso i pixel P_l e P_k del contorno corrispondenti all'estremità dell'asse minore.

1.3.7 Rettangolo di base

Il parallelogramma con lati paralleli a D_1 e D_2 che include completamente l'oggetto è chiamato rettangolo di base. Il rettangolo di base corrisponde al più piccolo rettangolo che include completamente l'oggetto e presenta una orientazione coincidente con quella dell'oggetto stesso.

Il rettangolo di base può essere calcolato se è noto l'angolo Θ di orientazione dell'oggetto.

Le coordinate (x,y) di ogni pixel P del contorno sono trasformate dalle equazioni di trasformazione geometrica per la rotazione, date da:

$$x' = x \cos\Theta + y \sin \Theta$$

$$y' = x \sin\Theta + y \cos \Theta$$

Dalle coordinate trasformate sono selezionate quelle con valore minimo e massimo rispettivamente dell'asse x' e y' individuando in questo modo sul contorno i corrispondenti quattro pixel P_1, P_2, P_3, P_4 che generano il rettangolo di base con base

$$a = (x'_{max}, x'_{min}) \text{ e altezza } h = (y'_{max}, y'_{min})$$

1.3.8 Eccentricità

La più semplice misura di eccentricità di un oggetto è ottenuta dal rapporto di D_1/D_2 tra le lunghezze dell'asse maggiore e quella dell'asse minore.

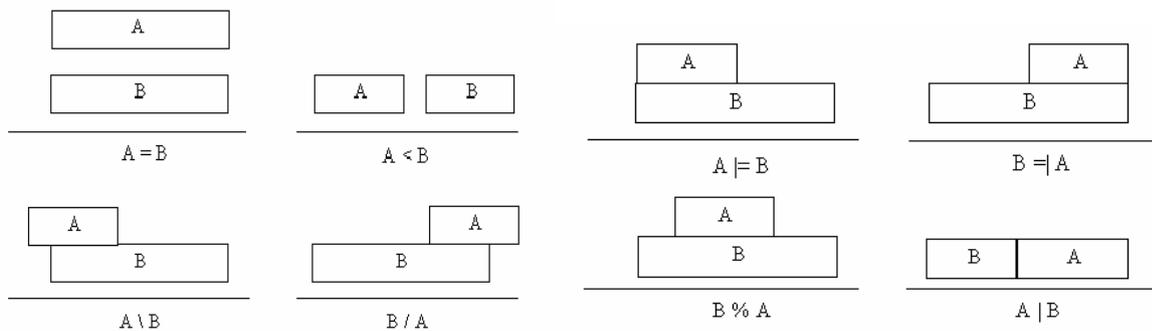
1.3.9 Allungamento

Una misura dell'allungamento di un oggetto è ottenuta dal rapporto tra l'altezza e la base del rettangolo di base associato allo stesso oggetto.

1.4 Informazioni spaziali

Regioni o oggetti con colori e proprietà di texture simili possono essere facilmente distinti imponendo dei vincoli spaziali. Per esempio una regione dell'immagine occupata dal cielo blu e una occupata dall'oceano potranno avere istogrammi di colore simili ma la loro

collocazione spaziale nell'immagine è differente. Una tecnica molto usata nella rappresentazione di relazioni spaziali tra oggetti è la 2D strings proposta da Chang et al [8]; è costruita proiettando le immagini lungo gli assi x e y. Definiamo quindi due insiemi di simboli V e A; ogni simbolo in V rappresenta un oggetto nell'immagine e ogni simbolo in A rappresenta un tipo di relazione spaziale tra oggetti (ad esempio: a destra di, a sinistra di, sopra a, sotto a). Una sua variante è il 2D G-string che taglia tutti gli oggetti lungo il loro bordo ed estende le relazioni spaziali tra oggetti in due insiemi di operatori spaziali. Un insieme definisce le relazioni spaziali locali e l'altro le relazioni spaziali globali e indica che la proiezione di due oggetti è disgiunta, congiunta o localizzata alla stessa posizione. Tuttavia la ricerca di un'immagine basata sulle relazioni spaziali tra oggetti o regioni è ancora materia di ricerca in quanto è difficile e alcune volte impossibile produrre un algoritmo di segmentazione affidabile se non in ambiti molto ristretti.



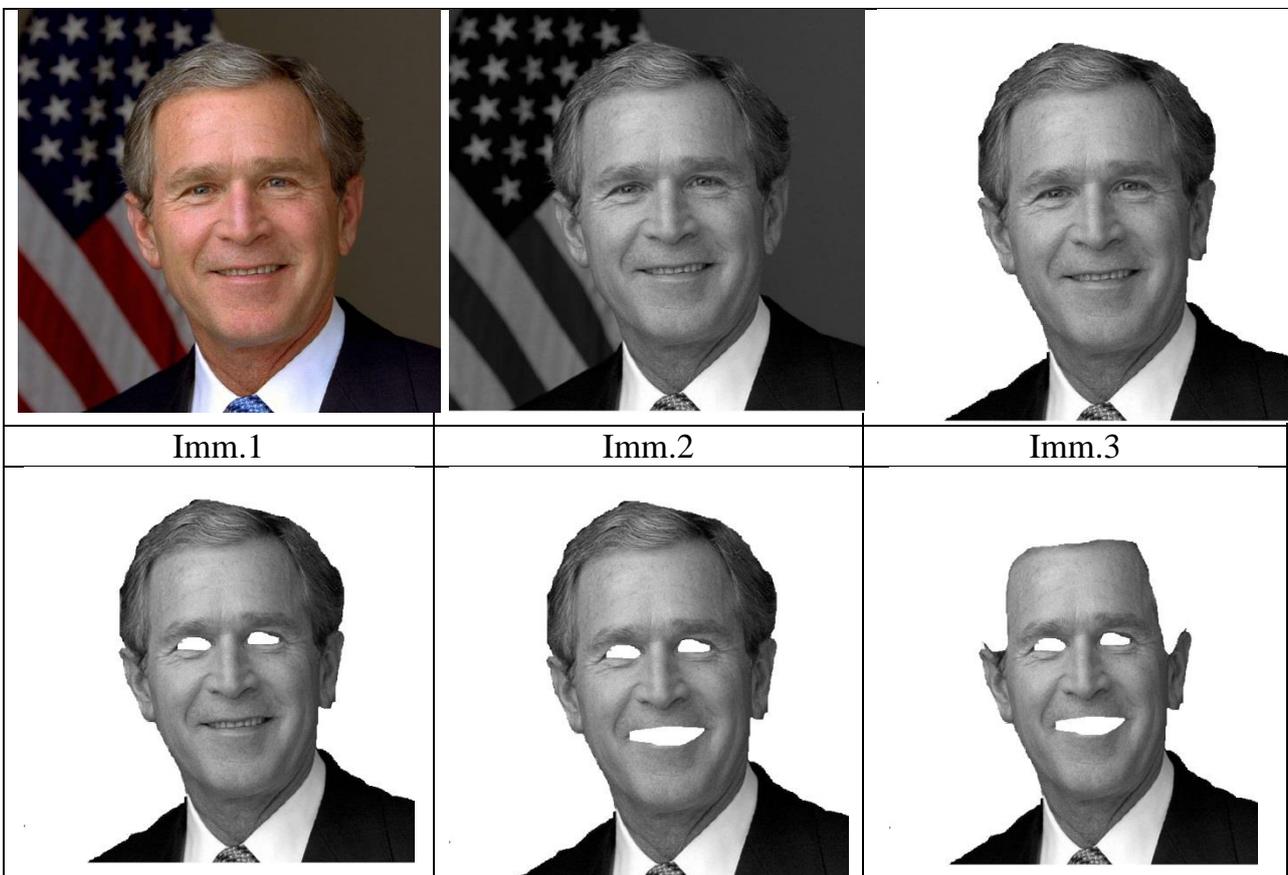
Notation	Condition	Meaning
$A < B$	$end(A) < begin(B)$	A disjoint B
$A = B$	$begin(A) = begin(B)$ and $end(A) = end(B)$	A equals B
$A B$	$end(A) = begin(B)$	A edge-to-edge B
$A \% B$	$begin(A) < begin(B)$ and $end(A) > end(B)$	A contains B (different bounds)
$A [B$	$begin(A) = begin(B)$ and $end(A) > end(B)$	A contains B (same begin)
$A] B$	$begin(A) < begin(B)$ and $end(A) = end(B)$	A contains B (same end)
A / B	$begin(A) < begin(B)$ and $end(A) < end(B)$	A partly -overlap B

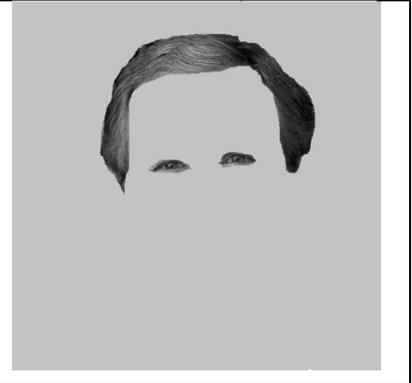
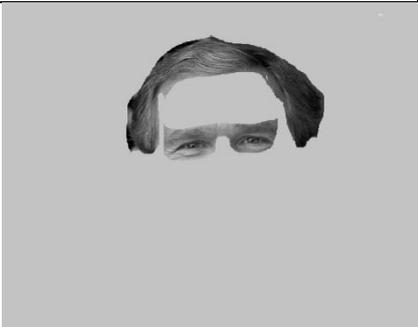
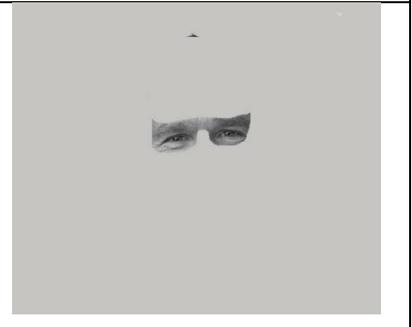
Fig.6 Esempio metodo 2D G-string.

Capitolo 2

Motori di ricerca: Google, TinEye e Yandex

In questo capitolo si descriveranno le interfacce dei tre motori di ricerca osservati, si cercherà di capire come Google classifichi un'immagine "visivamente simile" ad un'altra, e si confronteranno i motori di ricerca utilizzando alcune immagini test. Esse sono state ricavate modificando opportunamente un'immagine a colori di George W. Bush, le modifiche apportate all'immagine sono: rotazioni, diminuzioni di scala e perdita dell'informazione cromatica, perdita di particolari come lo sfondo, occhi, capelli etc. La scelta non è stata casuale, è stato scelto George W. Bush perché è un personaggio pubblico e il web è ricco di sue immagini modificate e non. Le immagini test sono state sottoposte ripetutamente ad ogni motore di ricerca per vedere se i risultati variavano nel numero e nel contenuto. Si è visto che ciò è una particolarità di Google e che questa caratteristica è indipendente dai cookie. Le ricerche sono state fatte utilizzando finestre anonime in quanto Yandex utilizza i cookie.



Imm.4	Imm.5	Imm.6
		
Imm.7	Imm.8	Imm.9
		
Imm.10	Imm.11	Imm.12
		
Imm.13	Imm.14	Imm.15
		
Imm.16	Imm.17	Imm.18

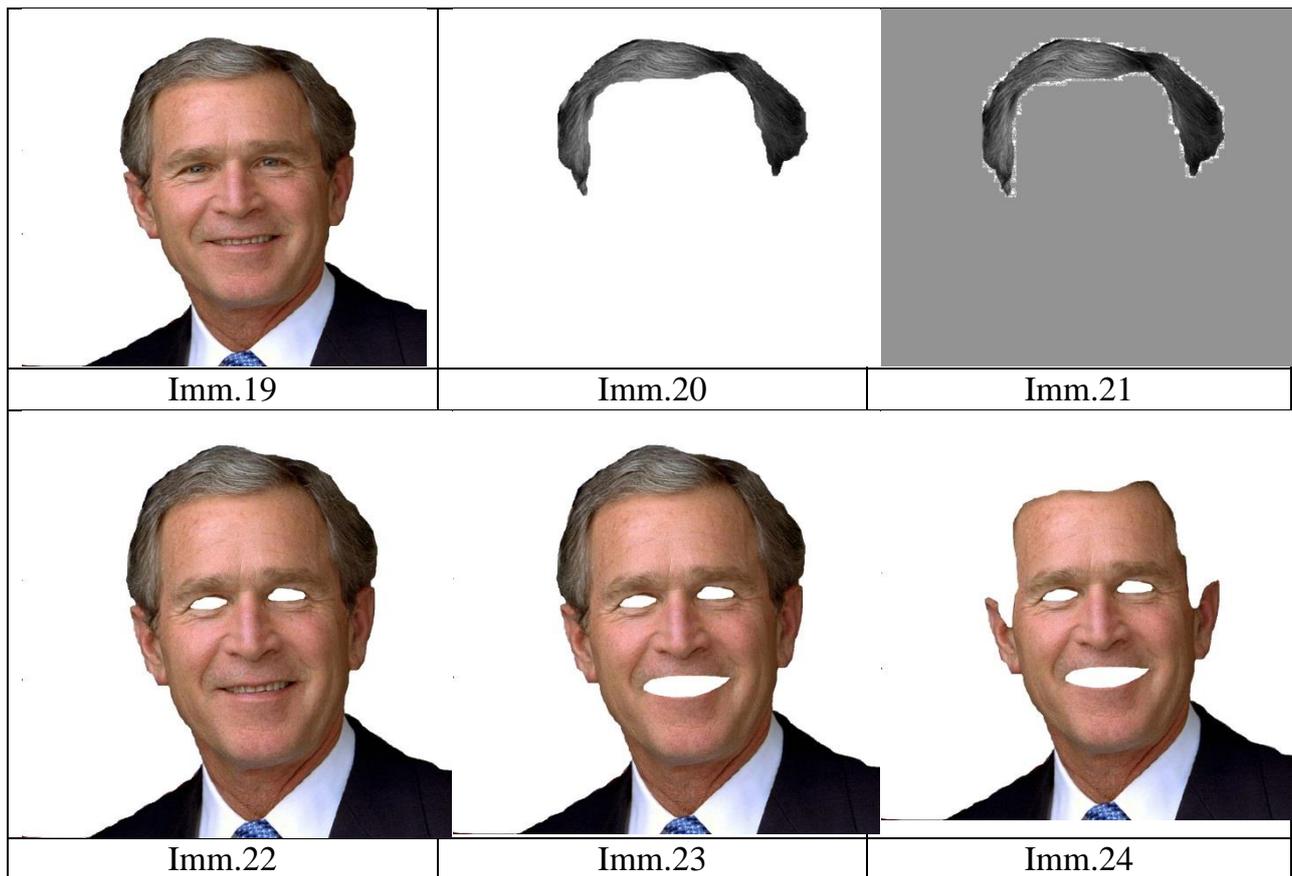


Fig.7 Immagini test.

2.1 Google

2.1.1 Interfaccia utente

L'interfaccia che Google presenta all'utente è molto facile da usare, basta cliccare 'immagini' in alto a destra per selezionare il motore di ricerca per immagini. A questo punto si può trascinare l'immagine sulla barra di ricerca posta al centro. Il motore di ricerca indicherà il numero di risultati, il tempo impiegato nella ricerca, la dimensione dell'immagine usata, se si vuole cercare la stessa immagine con altre dimensioni, le immagini visivamente simili e i siti in cui si può trovare l'immagine test, originale o modificata.



Fig.8 Esempio di ricerca con Google.

Cliccando su ‘tutte le dimensioni’ si possono filtrare i risultati ottenuti per colore, dimensione, diritti di utilizzo e data.

2.1.2 Immagini visivamente simili

Google offre un servizio che nessun altro motore di ricerca offre: cerca immagini visivamente simili a quella sottoposta. Per capire cosa intenda Google per immagini “visivamente simili”, si è sottoposto il motore di ricerca ad alcune immagini test qui di seguito elencate.

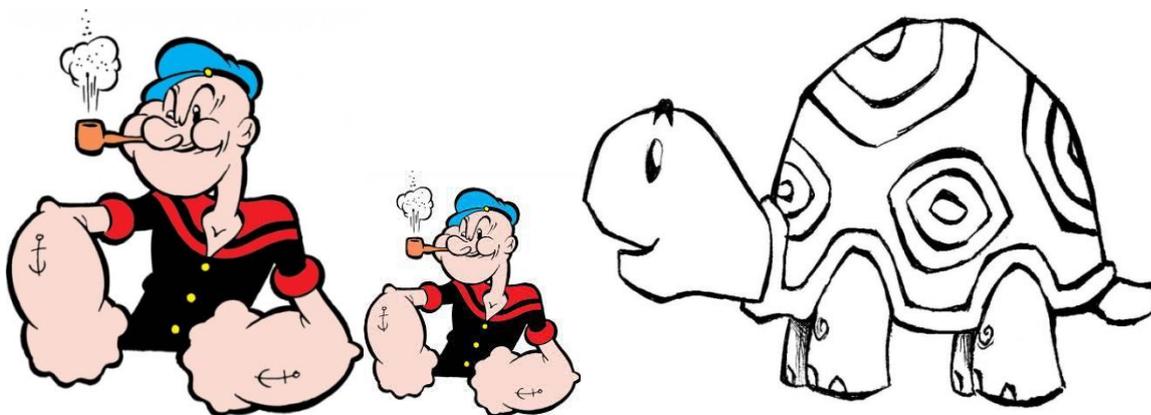


Fig.9 Immagini con differente scala e istogrammi a colori e scale di grigio.



Fig.10 Immagini con differente direzionalità, rispettivamente casuale, obliqua, orizzontale, verticale.



Fig.11 Immagini con texture a granularità fine e grossolana.



Fig.12 Immagine con texture regolare.

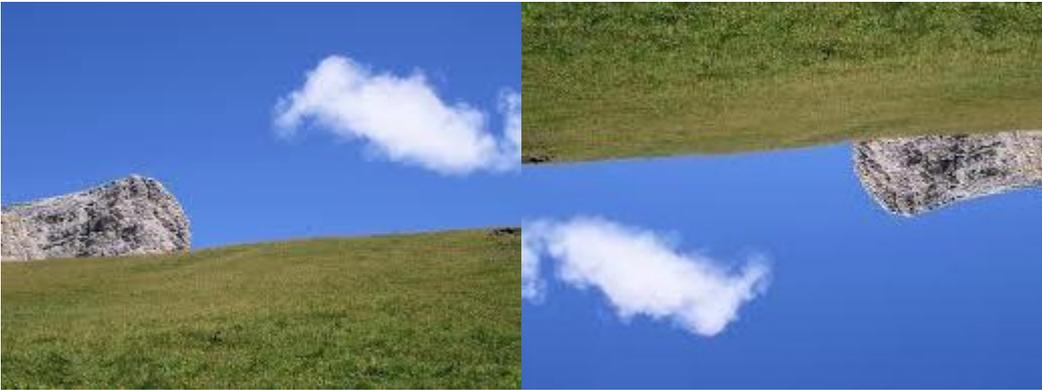


Fig.13 Immagini utilizzate per osservare se Google riconosce le relazioni spaziali.

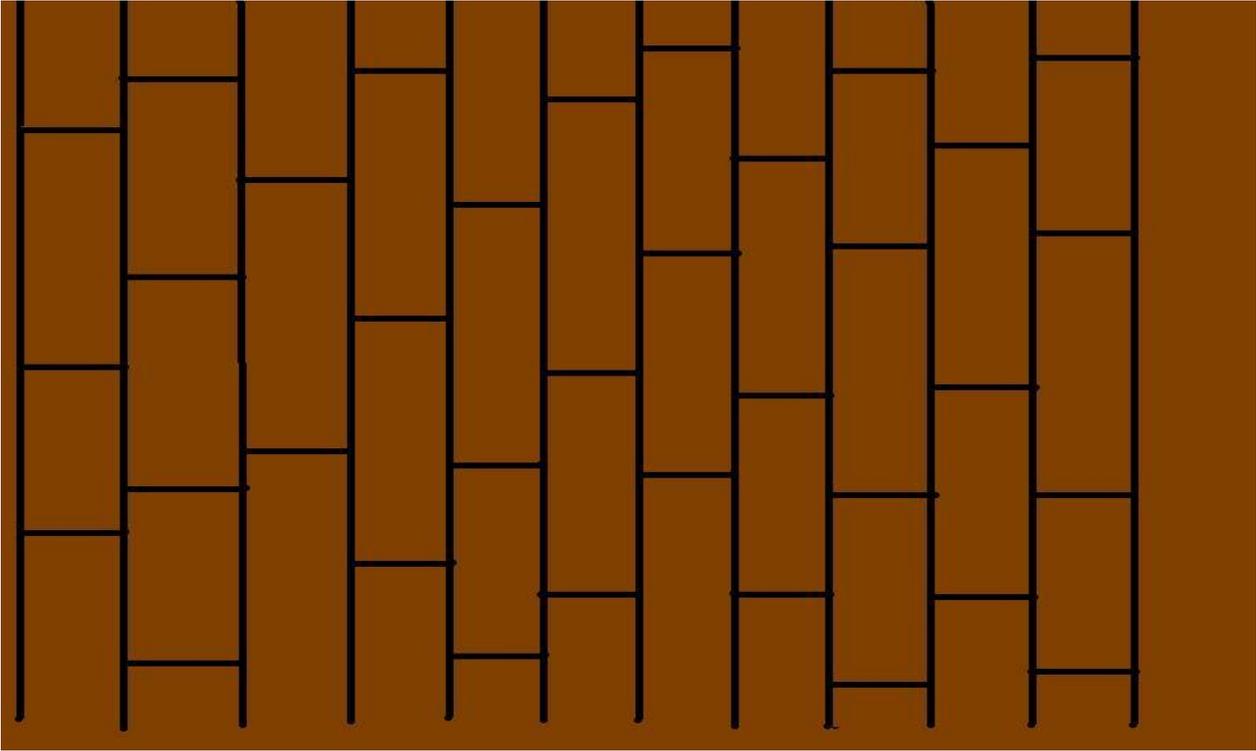


Fig.14 Immagine utilizzata per osservare se Google associa all'immagine caratteristiche locali.



Fig.15 Immagine di George Clooney

2.1.2.1 Risultati

Sottoponendo le immagini in Fig.9 al motore di ricerca Google, si osserva che il sistema associa una descrizione testuale all'immagine coerente con la lingua del paese in cui ci si è collegati. Nel nostro caso: Braccio di Ferro. La descrizione testuale è molto importante perché restringe di molto il campo di ricerca.



Fig.16 Risultato ottenuto da Google per l'immagine di Braccio di Ferro.

Nelle immagini visivamente simili si trovano principalmente immagini a colori, molte delle quali con istogrammi simili. Quelle in bianco e nero o con istogrammi differenti hanno nella descrizione testuale la parola chiave Braccio di Ferro, come i personaggi del fumetto, e il logo della lega di braccio di ferro.



Fig.17 Immagini ‘visivamente simili’ a quella di Braccio di Ferro.

Ragionare per parole chiave è conveniente specialmente quando si cerca un elemento di un dato insieme di immagini e si possiede l’immagine non del soggetto che interessa ma di uno ad esso correlato. Occorre tuttavia classificare le immagini in modo opportuno. Un esempio in cui la classificazione, a mio avviso, fallisce è quando si sottopone al sistema l’immagine della tartaruga. La parola chiave che esso le associa è: disegni da colorare. Abbiamo come risultato tutte immagini in bianco e nero ma semanticamente differenti.



Fig.18 Risultato fornito da Google per l’immagine della tartaruga.

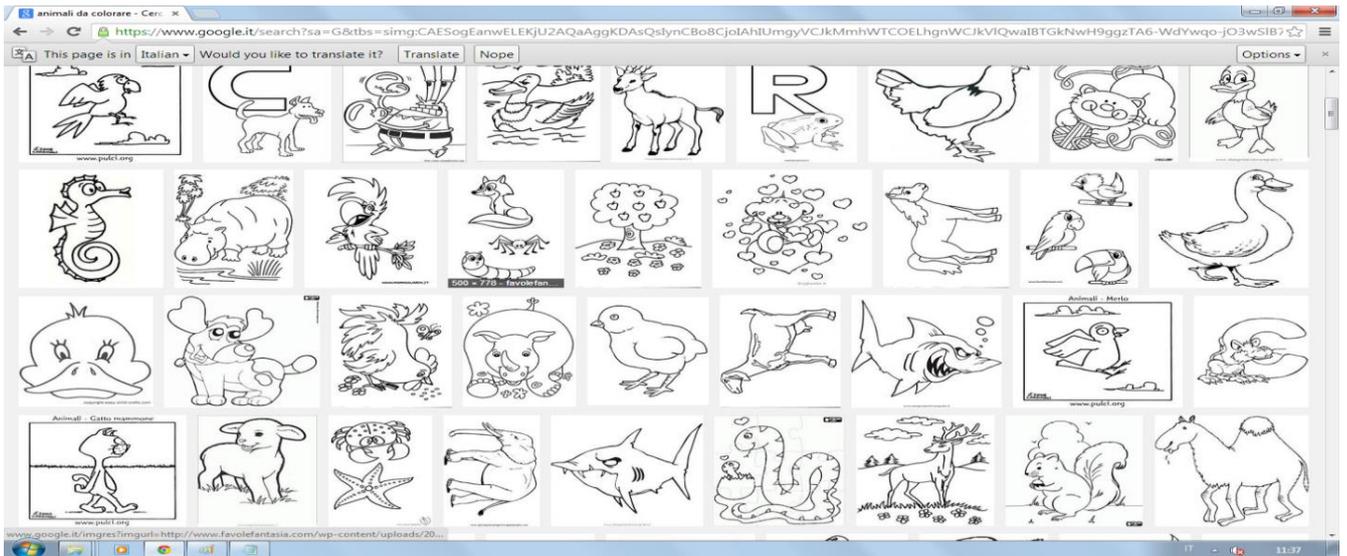


Fig.19 Immagini 'visivamente simili' a quella della tartaruga.

Sottoponendo un'immagine con scala diversa in cui le proporzioni siano rispettate si osserva che le immagini visivamente simili sono le stesse e si può supporre che non solo viene utilizzato l'istogramma delle immagini ma che è normalizzato per il numero dei pixel.

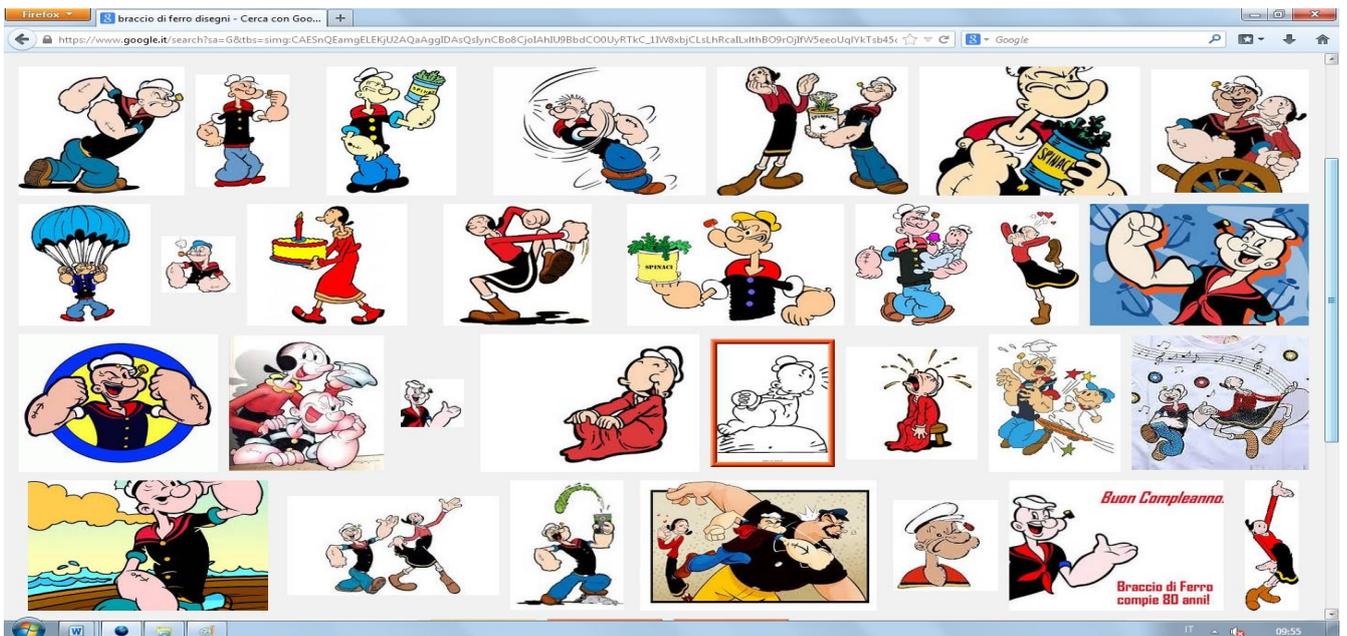


Fig.20 Risultato per l'immagine di Braccio di ferro 524x600 pixel.

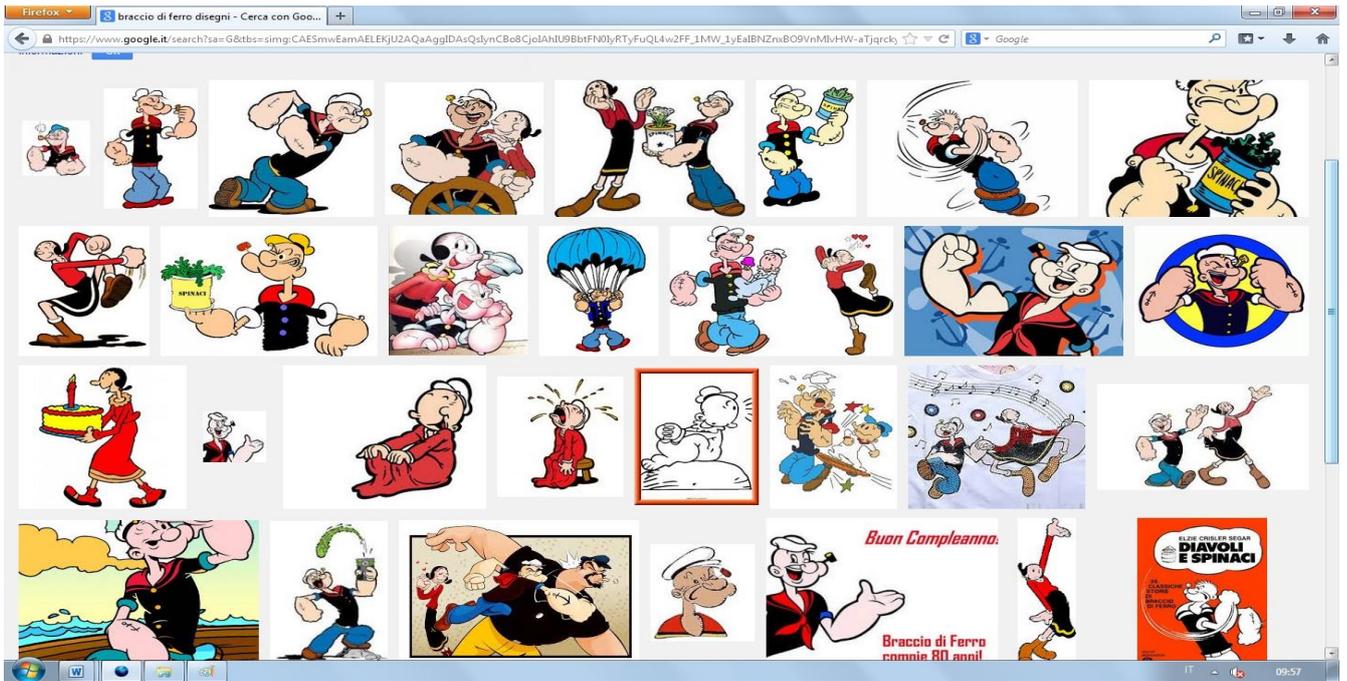


Fig.21 Risultato per l'immagine di braccio di ferro 210x240 pixel.

Osservando i risultati per le immagini di Fig.10 si può osservare che il sistema cerca immagini con la stessa direzionalità e istogramma.

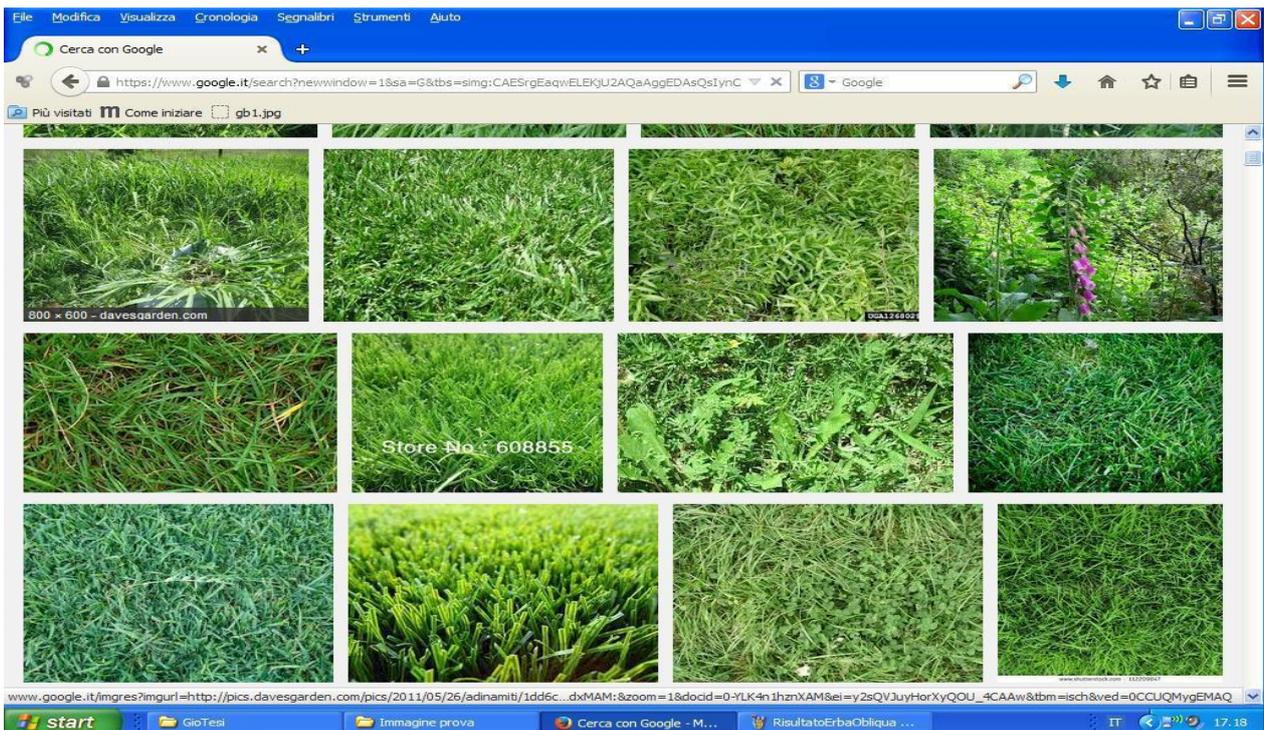


Fig.22 Risultato nel caso di direzionalità casuale.

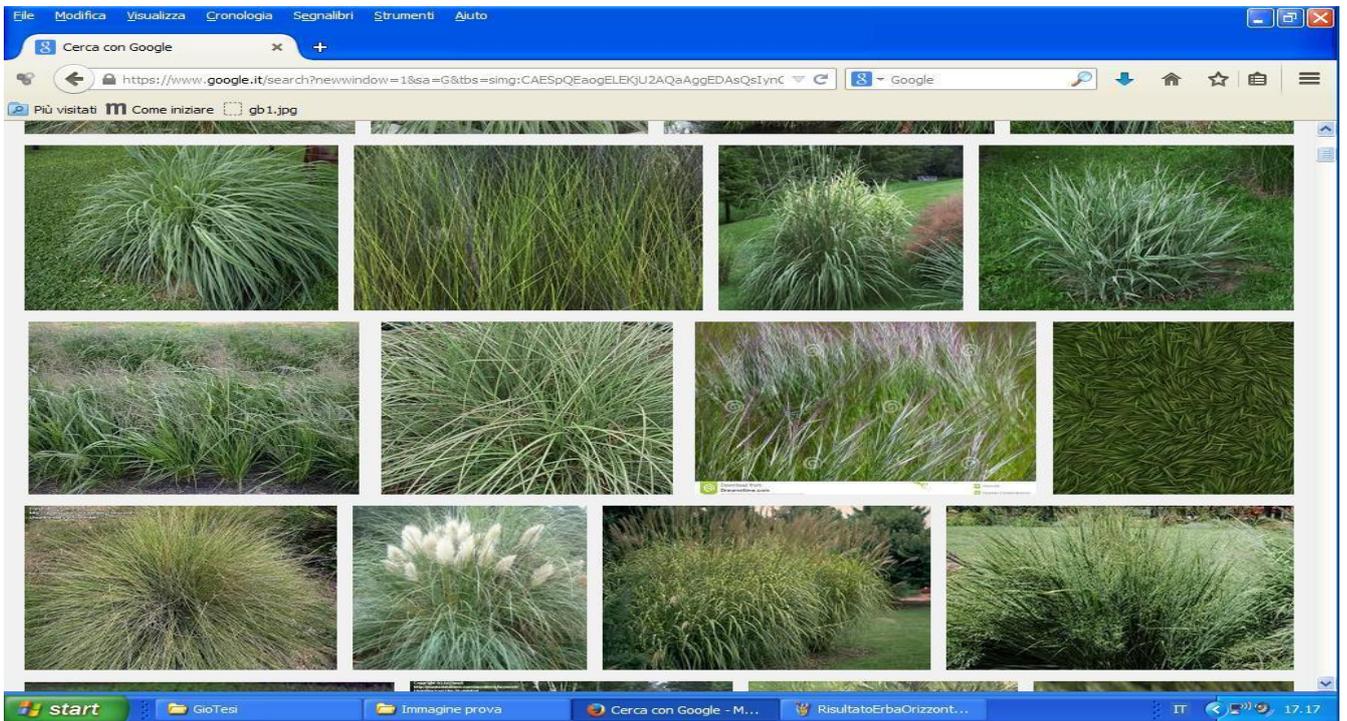


Fig.23 Risultato nel caso di direzionalità obliqua.

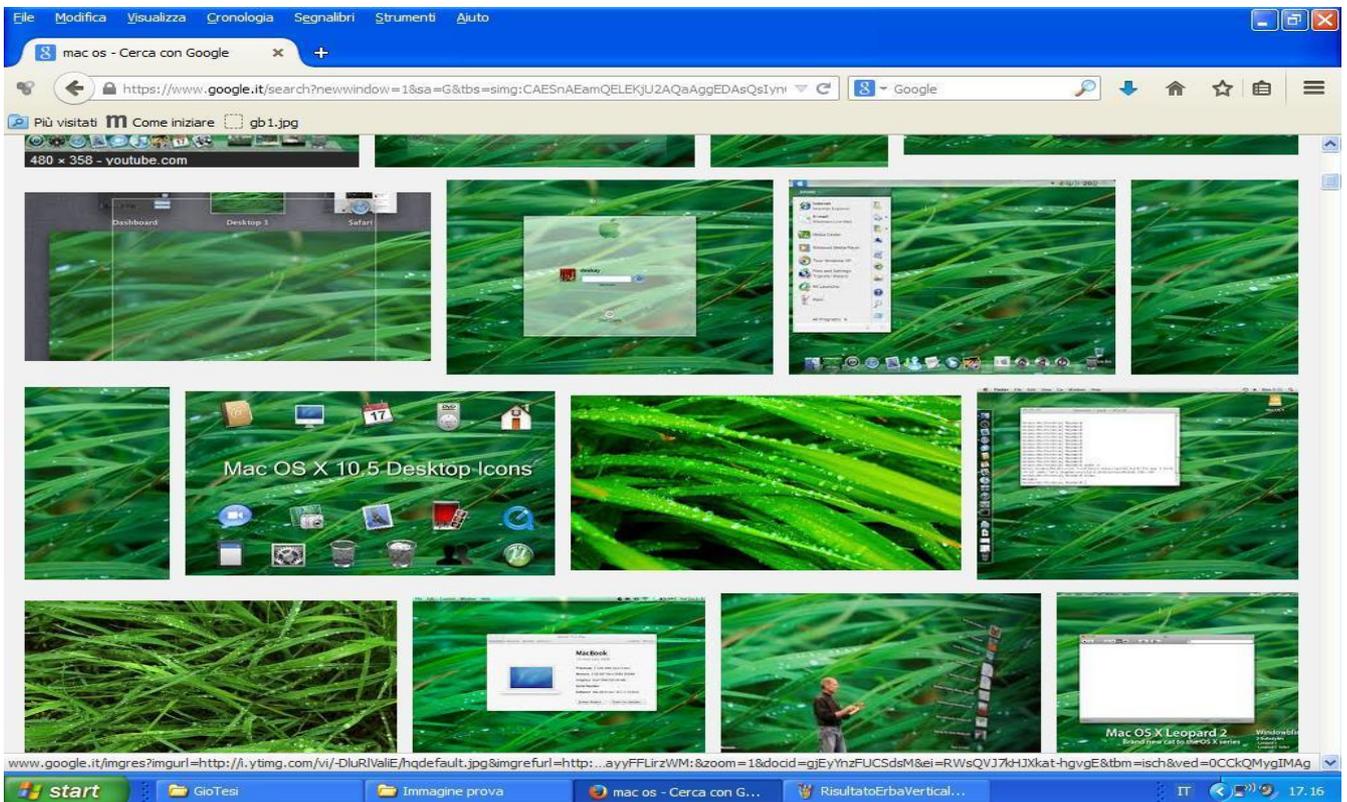


Fig.24 Risultato nel caso di direzionalità orizzontale.

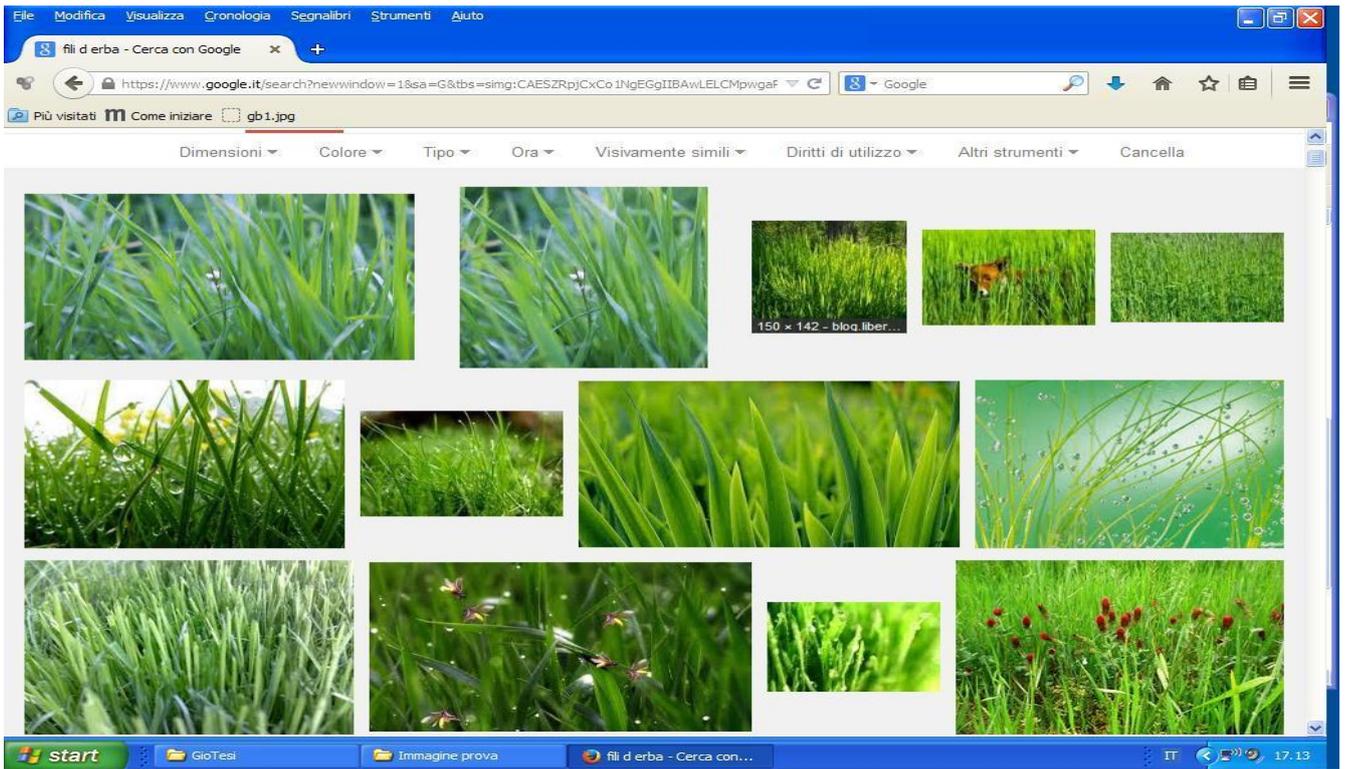


Fig.25 Risultato nel caso di direzionalità verticale.

I risultati per le Fig.11 mostrano come il sistema riconosca i diversi tipi di texture discriminando le diverse granularità in questo caso fine e grossolana.

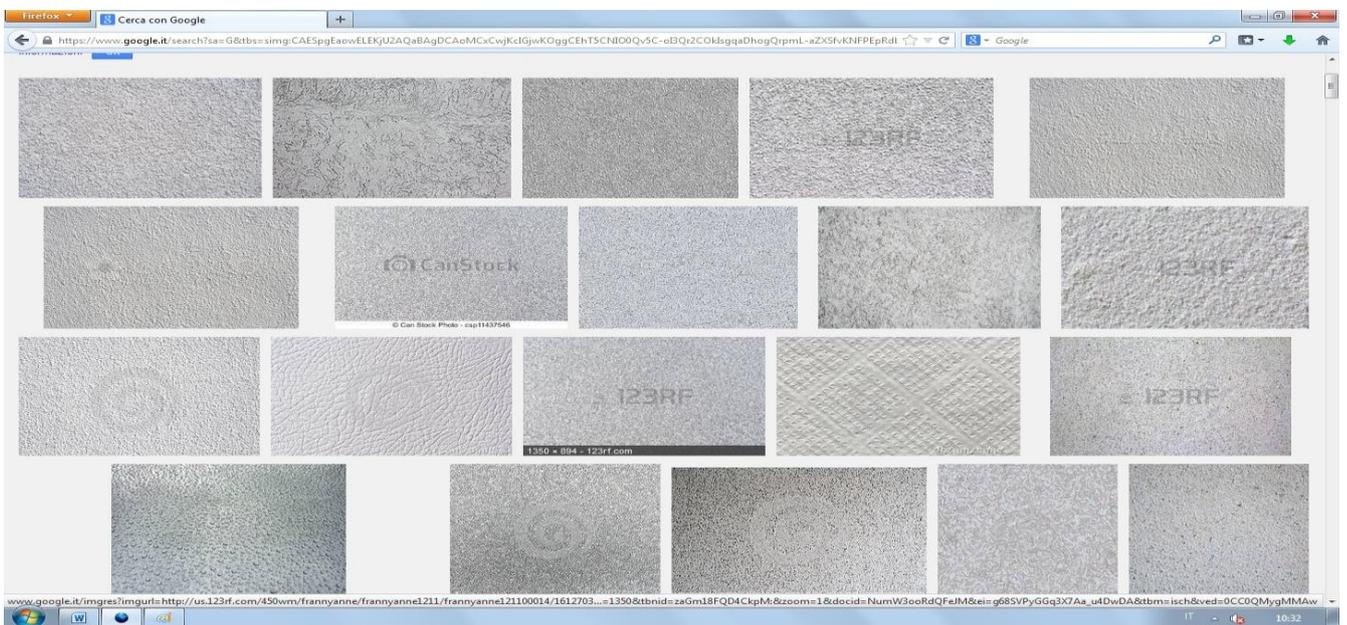


Fig.26 Risultato granularità fine.

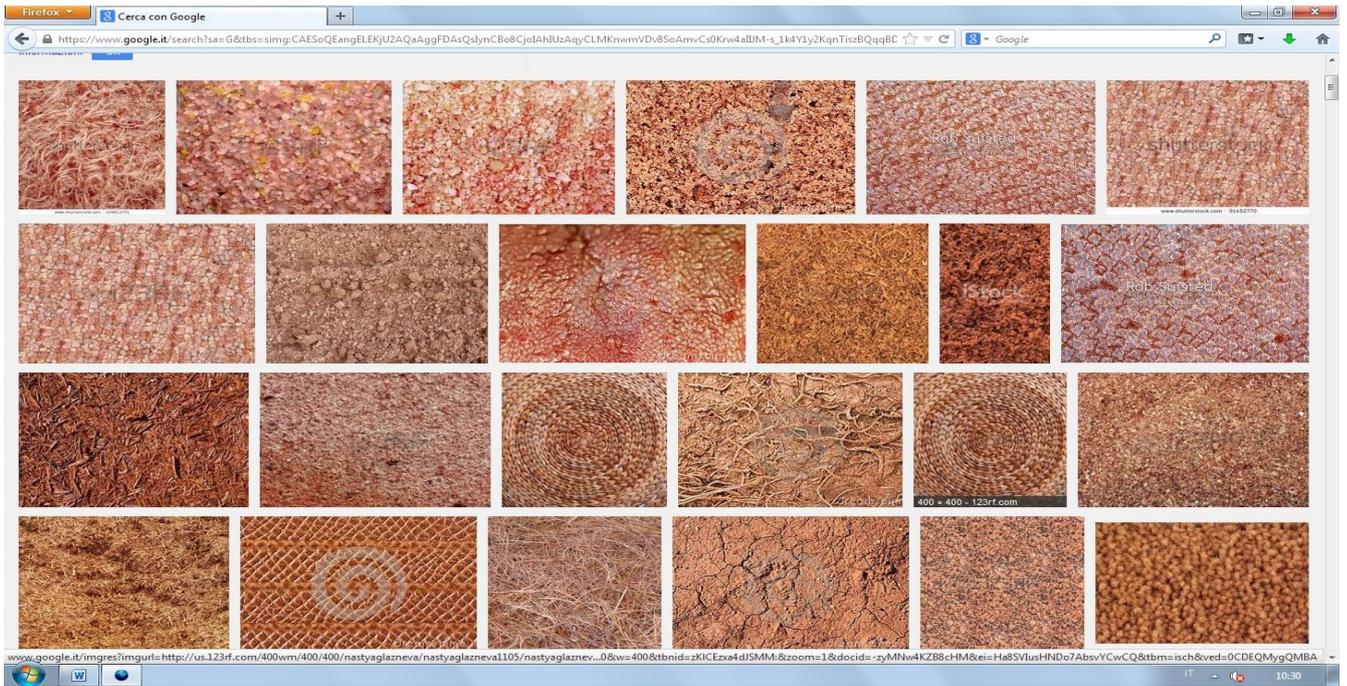


Fig.27 Risultato granularità grossolana.

Il sistema risponde bene anche per immagini con texture regolari come in Fig.12.

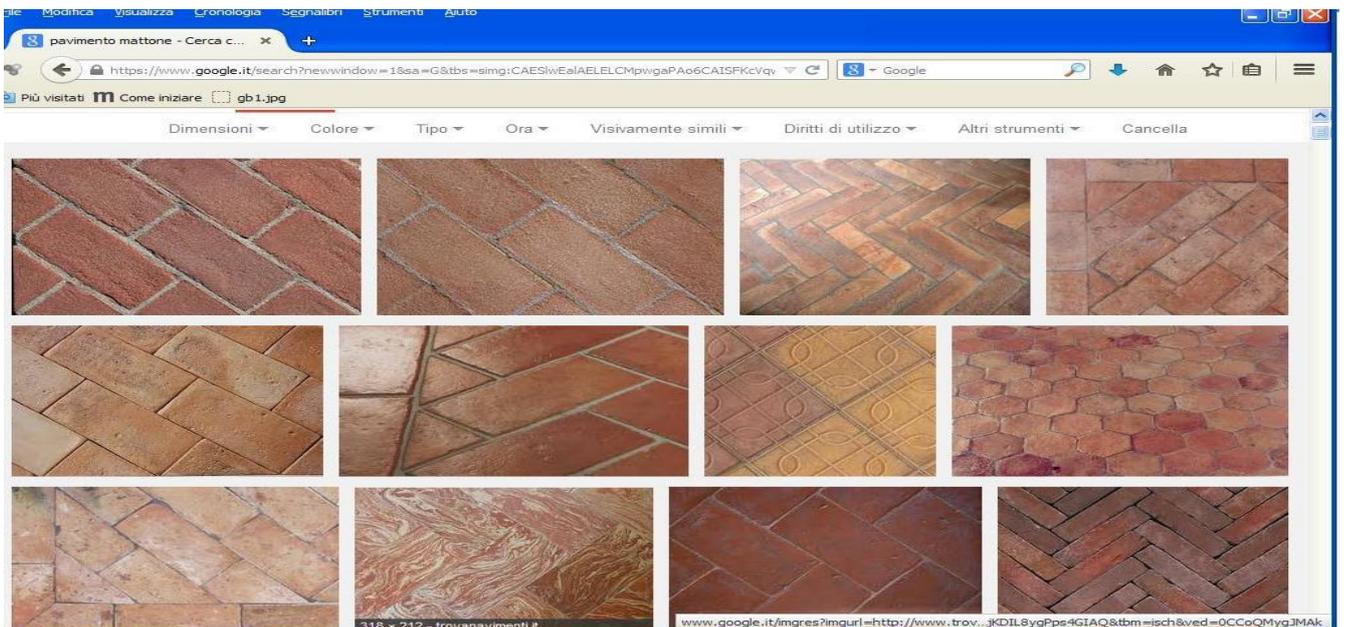


Fig.28 Risultato nel caso di texture regolare.

Sottoponendo le immagini in Fig.13 si può supporre che Google utilizzi algoritmi che riconoscono le relazioni spaziali: cielo e mare possono avere istogrammi simili ma la loro posizione nell'immagine è differente. Si è scelta un'immagine con istogramma simile a quella che può avere un'immagine con mare e terra e la si è rotata di 180°. I risultati concordano con la supposizione appena fatta.



Fig.29 Risultato ottenuto con l'immagine non ruotata.

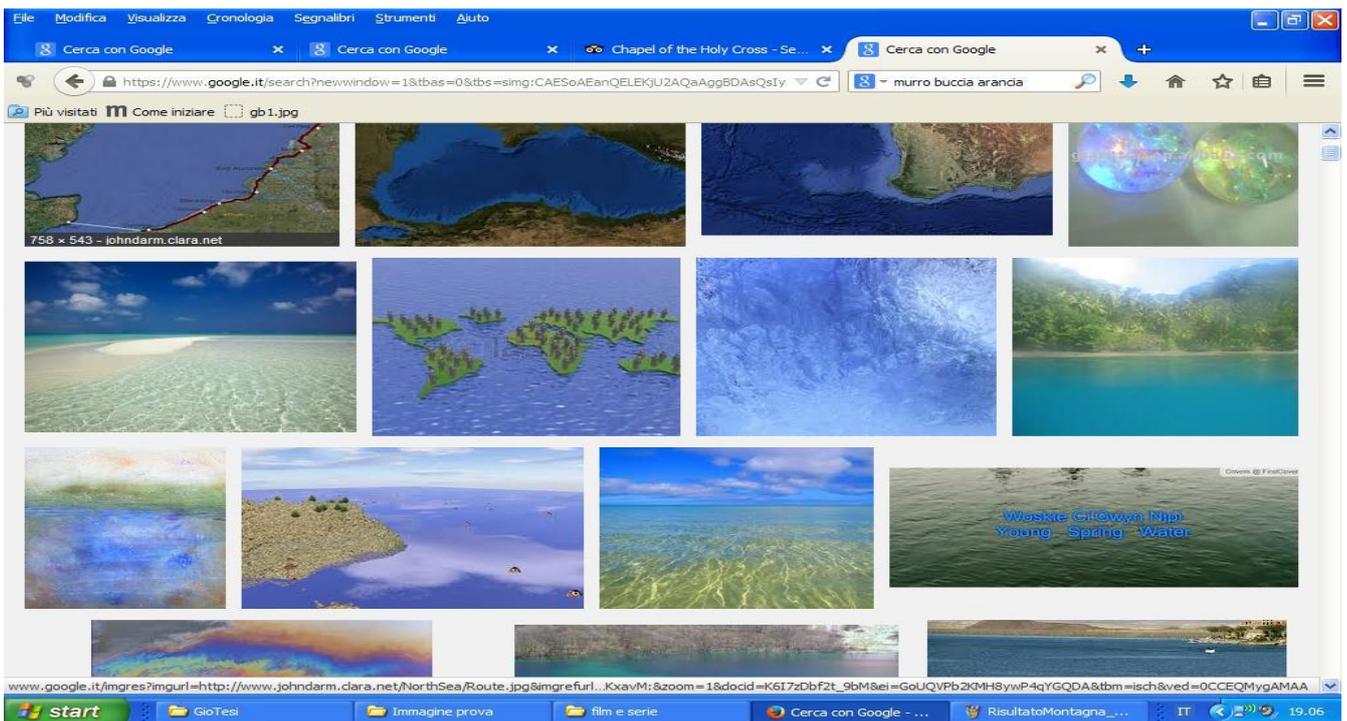


Fig.30 Risultato ottenuto con l'immagine rotata.

Sottoponendo la Fig.14 a Google, la associa a una chiesa particolare.

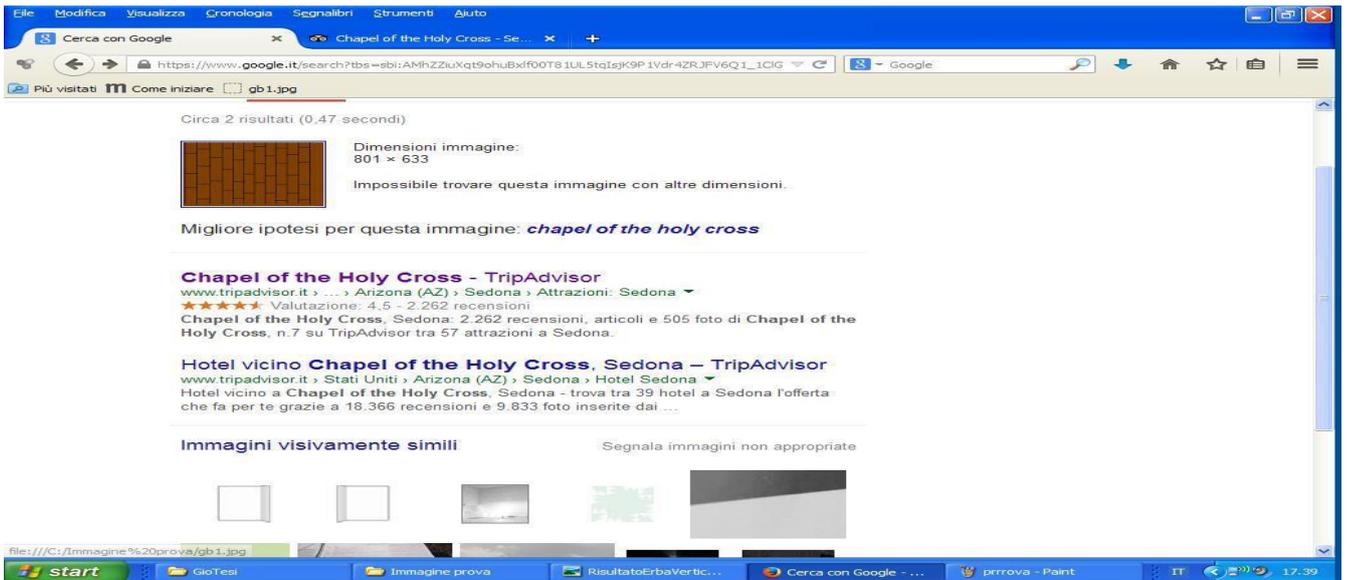


Fig.31 Risultato per l'immagine in Fig.14.

Cliccando sul link suggerito da Google otteniamo una serie di immagini. Si è scelta la più significativa.

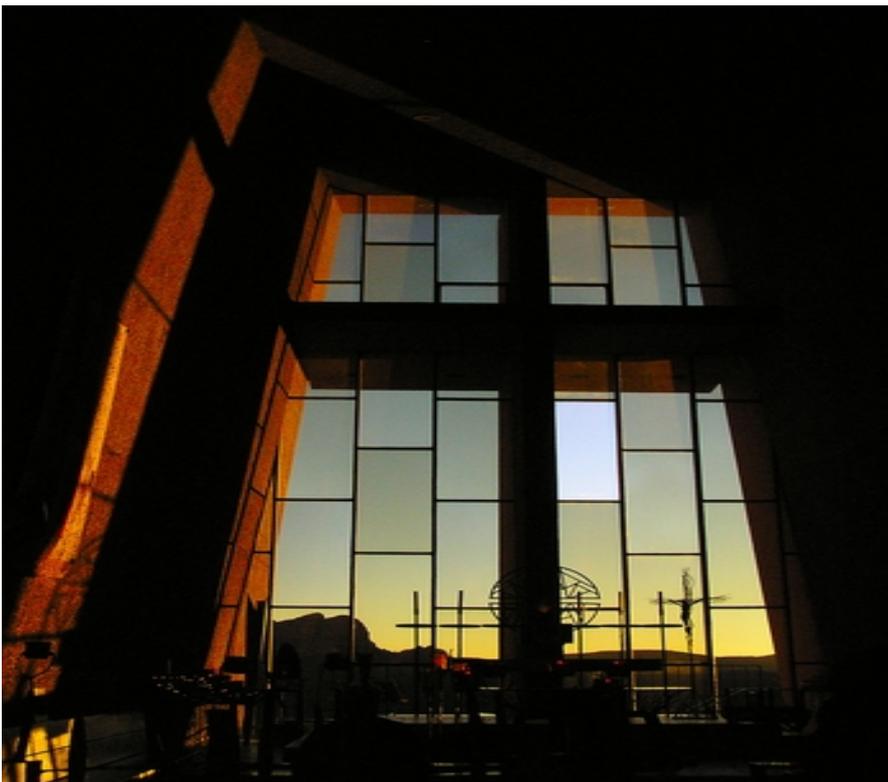


Fig.32 vetrata absidale della chiesa.

Questa immagine presenta localmente una texture con la stessa regolarità dell'immagine test.

Sottoponendo la Fig.15 e ripetendo le ricerche più volte si può osservare che non solo il numero dei risultati varia ma varia anche l'ipotesi che fa il sistema sul riconoscimento dell'immagine.

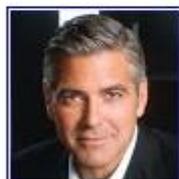
Firefox Cerca con Google Abside - Wikipedia

https://www.google.it/search?tbs=sbi:AMhZZivtd6CKIXZlwu9ZR0BEWx4kCfl9OJMm3DeH8KxpPfpBKUf

Google A1.jpg x george clooney

Web **Immagini** Notizie Shopping Maps Altro ▾ Strum

Circa 209 risultati (0,71 secondi)



Dimensioni immagine:
800 × 816

Trova altre dimensioni di questa immagine:
Tutte le dimensioni - [Medie](#)

Migliore ipotesi per questa immagine: **george clooney**

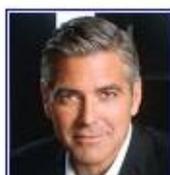
Firefox Cerca con Google Abside - Wikipedia

https://www.google.it/search?tbs=sbi:AMhZZitBnsve6P7sznKY9EKd8HadM8th-AiNACdDhNOirCJ49trpbYmpQzedKZqBl

Google A1.jpg x george clooney

Web **Immagini** Notizie Shopping Maps Altro ▾ Strumenti di ricerca

Circa 250 risultati (0,65 secondi)



Dimensioni immagine:
800 × 816

Trova altre dimensioni di questa immagine:
Tutte le dimensioni - [Medie](#)

Migliore ipotesi per questa immagine: **george clooney**

Fig.33 Risultato numero 1 e 2 fornito da Google per l'immagine di George Clooney in cui varia solo il numero dei risultati.



Fig.34 Risultato numero 2 fornito da Google per l'immagine di George Clooney in cui varia l'ipotesi fatta.

Si può supporre che Google utilizzi un certo numero di immagini che reputa simili all'immagine sottoposta e che le proponga come soluzione casualmente: è indipendente sia dal numero di volte che l'immagine test si propone sia dai cookie infatti le stesse soluzioni vengono proposte dal motore di ricerca usando una finestra anonima.

2.1.3 Risultati di Google per le immagini test in fig.7

Sottoponendo a Google le immagini in fig.7, si ottengo i risultati riportati nella seguente tabella.

Mot.Ricerca / Immagini	Google	
	Numero Risultati	Match

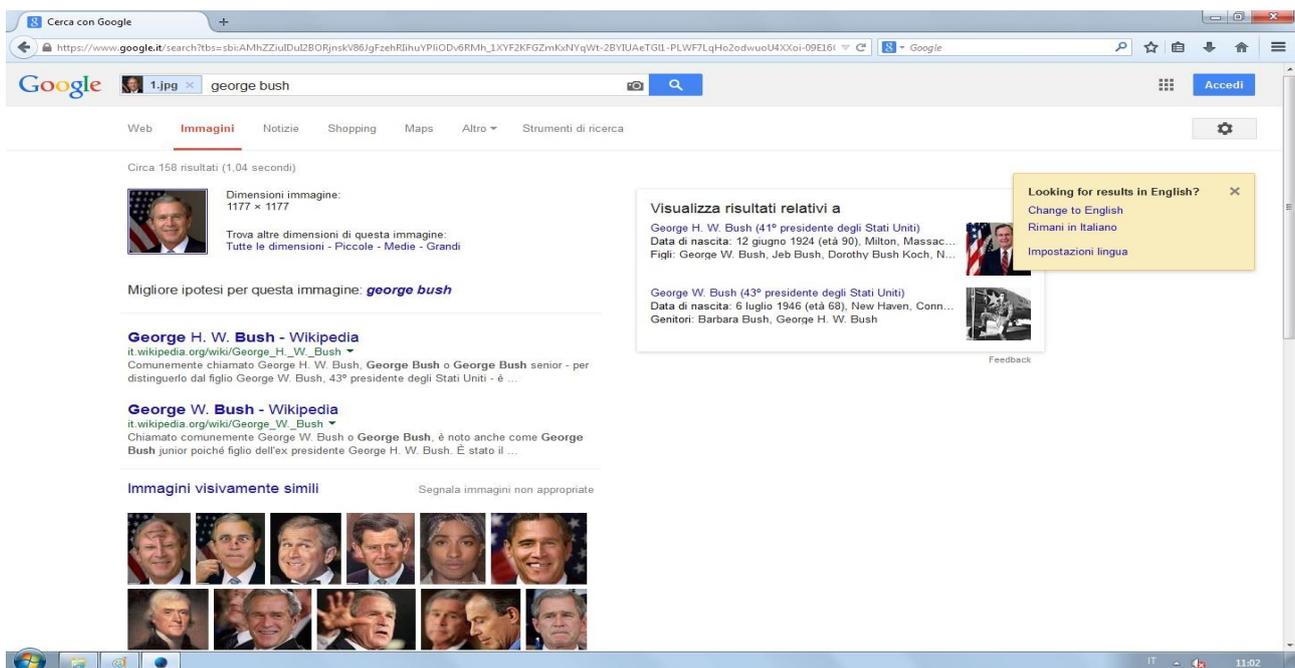
Imm.1	138	Sì
Imm.2	102	No
Imm.3	204	Sì*
Imm.4	61	No
Imm.5	35	No
Imm.6	20	No
Imm.7	1	Sì*
Imm.8	0	No
Imm.9	0	No
Imm.10	5	Sì*
Imm.11	4	Sì
Imm.12	2	Sì
Imm.13	0	No
Imm.14	28	Sì
Imm.15	55	Sì*
Imm.16	0	No
Imm.17	3	Sì
Imm.18	5	Sì
Imm.19	162	No
Imm.20	0	No

Imm.21	0	No
Imm.22	77	No
Imm.23	32	No
Imm.24	20	Sì

Fig.35 Risultati ottenuti da Google per le immagini in fig.7

Per Google il numero di risultati rappresenta il numero di siti in cui si trova l'immagine sottoposta a ricerca. Il 'sì' asteriscato nella tabella vuol dire che l'immagine campione sottoposta più volte al motore di ricerca ha dato risultati sia coerenti che non coerenti.

Immagine 1: Google associa all'immagine test più immagini, tutte pertinenti, presentandole random cambiando anche la parola chiave associata.



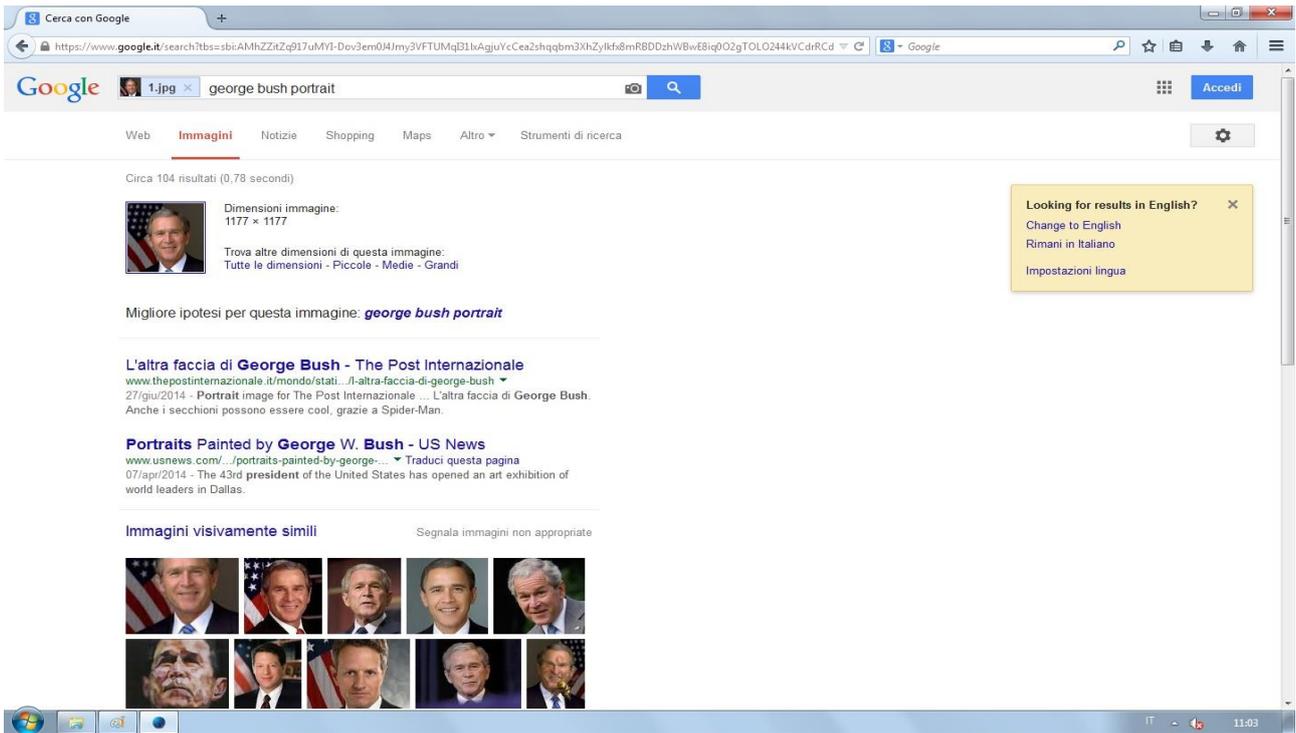


Fig.36 Risultati ottenuti sottoponendo l'immagine originale a Google.

Immagine 2: Google dà risultati ma non coerenti.

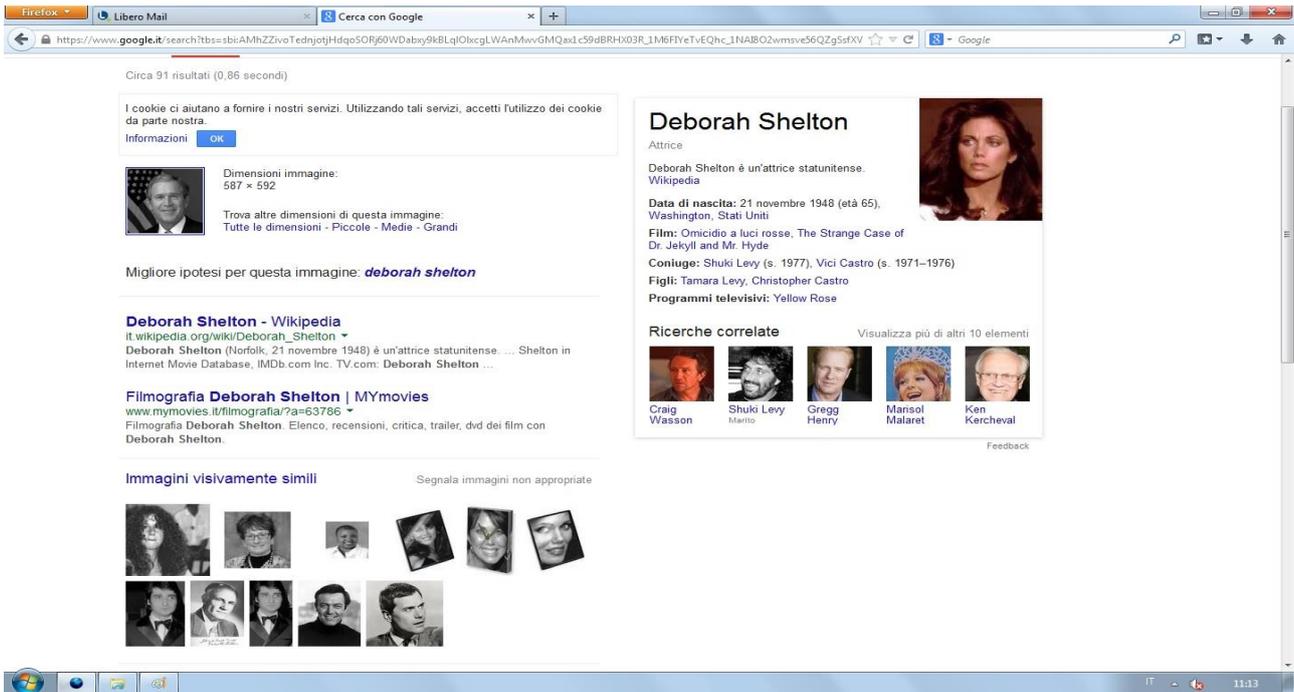


Fig.37 Risultato non coerente fornito da Google per l'immagine di G. Bush in bianco e nero.

Ciò è dovuto al fatto che Google non ha nel database un'immagine in bianco e nero di George Bush che si avvicina all'immagine test.

Immagine 3: Google ha due immagini che si avvicinano all'immagine test, una coerente e l'altra non coerente e le presenta random quando gli si sottopone l'immagine test.

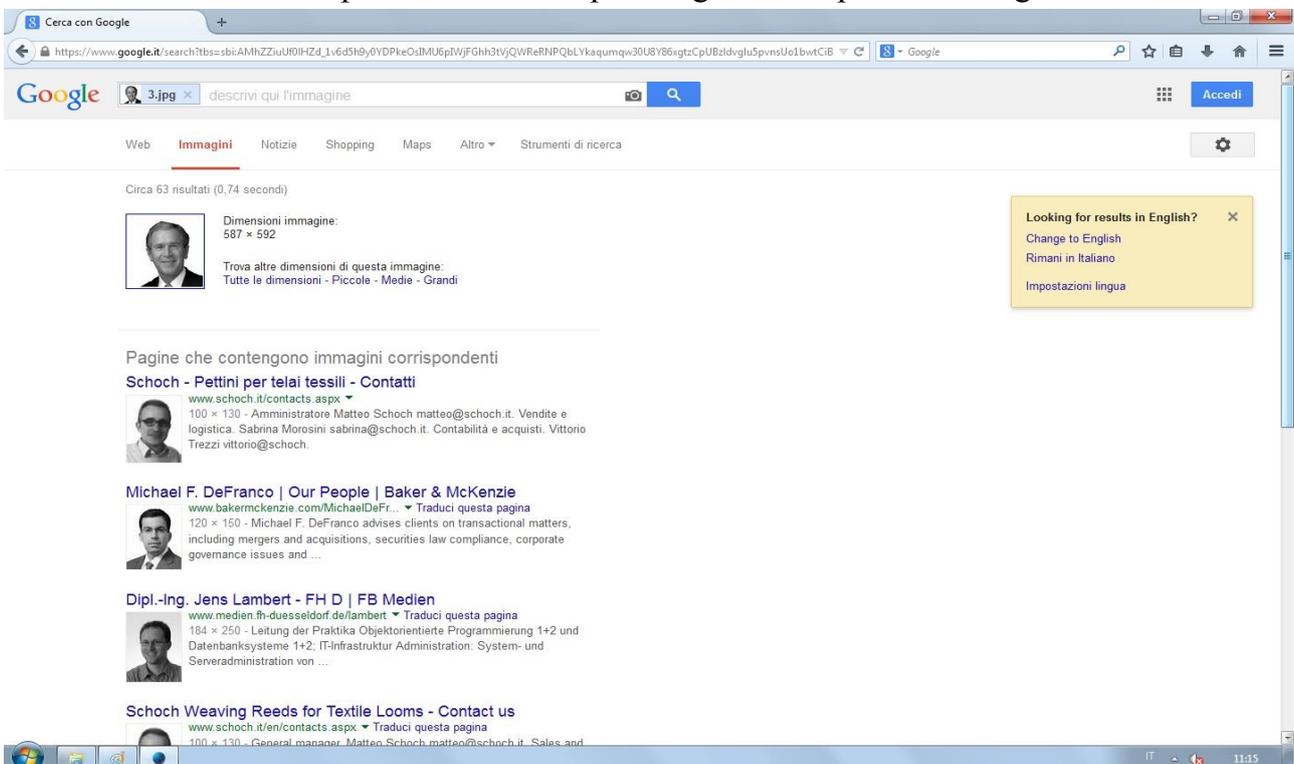


Fig.38 Risultato non coerente per l'immagine 3.

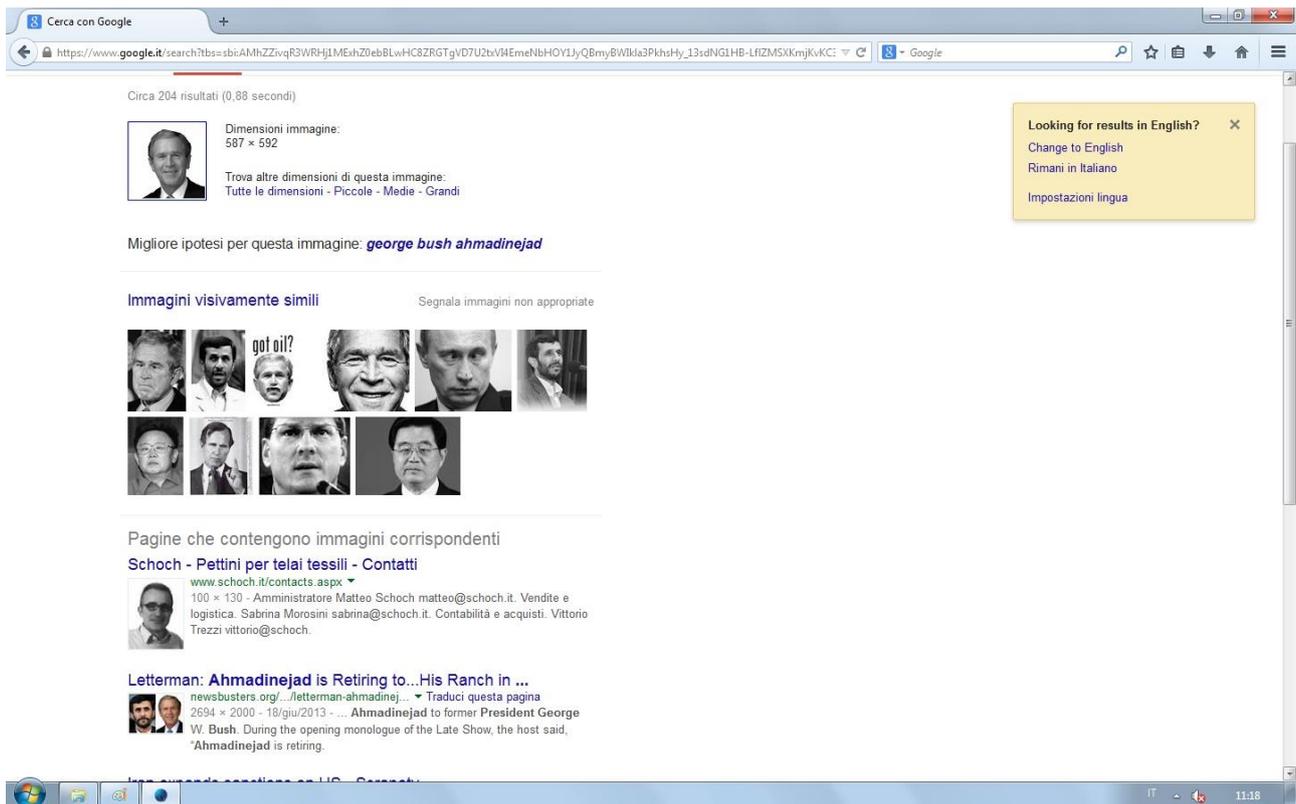


Fig.39 Risultato coerente per l'immagine 3.

Immagine 4, 5, 6: Google dà risultati non coerenti. Togliendo gli occhi l'immagine che veniva presentata random come risultato nell'immagine 3 non appare più.

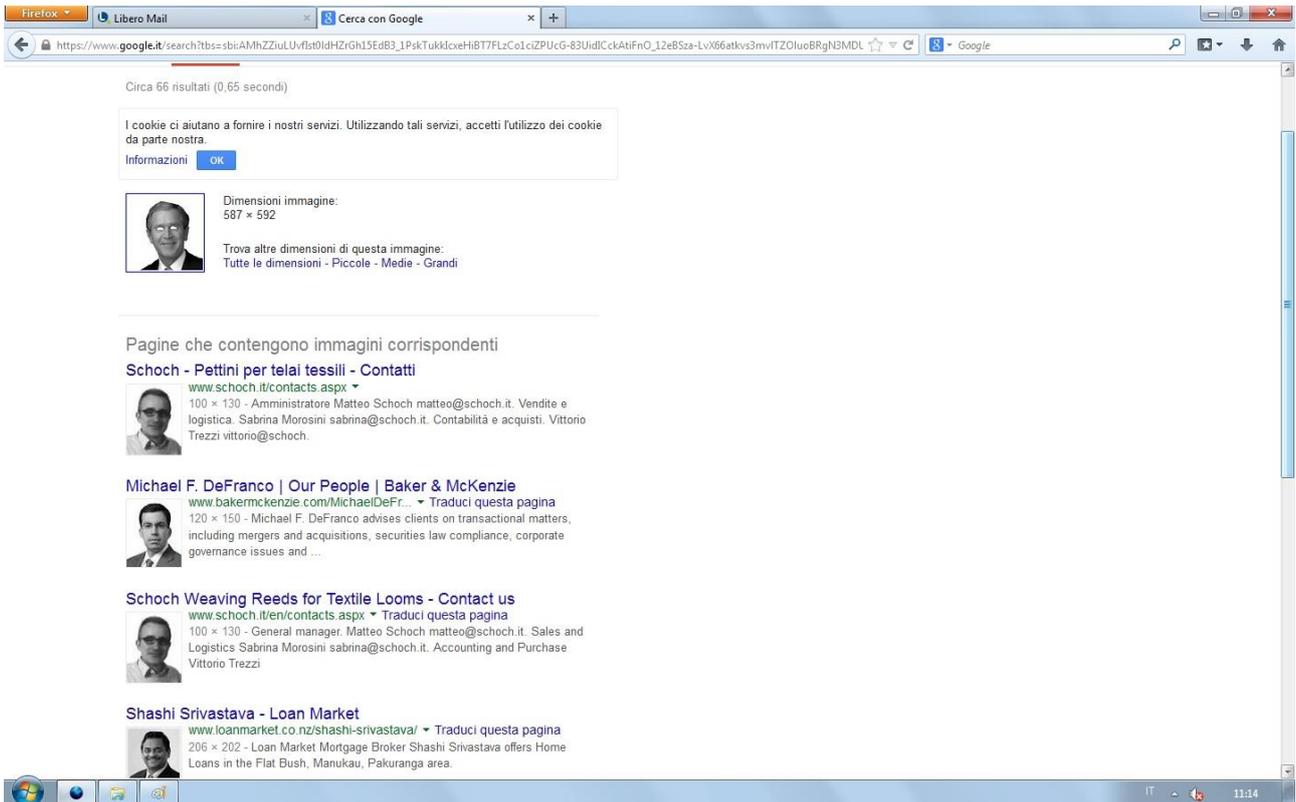


Fig.40 Risultati non coerenti ottenuti da Google per l'immagine 4.

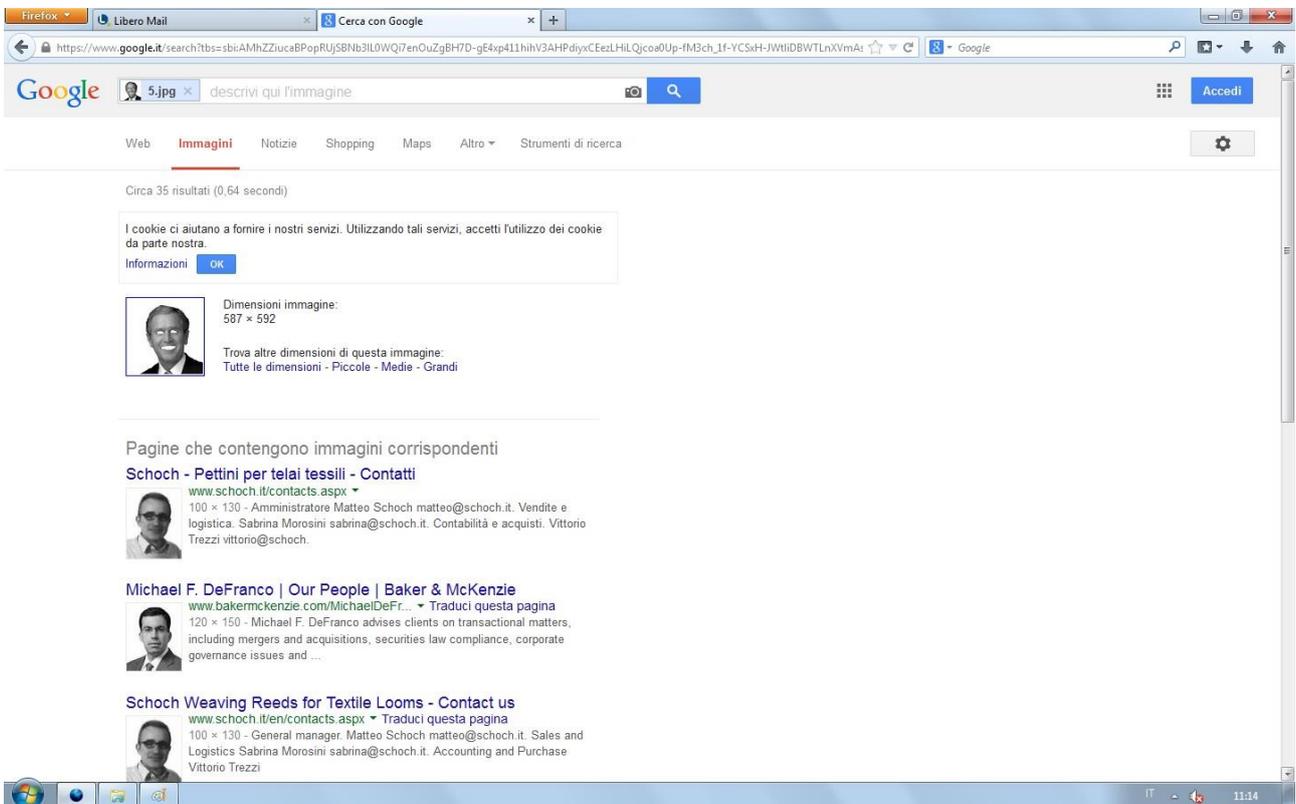


Fig.41 Risultati non coerenti ottenuti da Google per l'immagine 5.

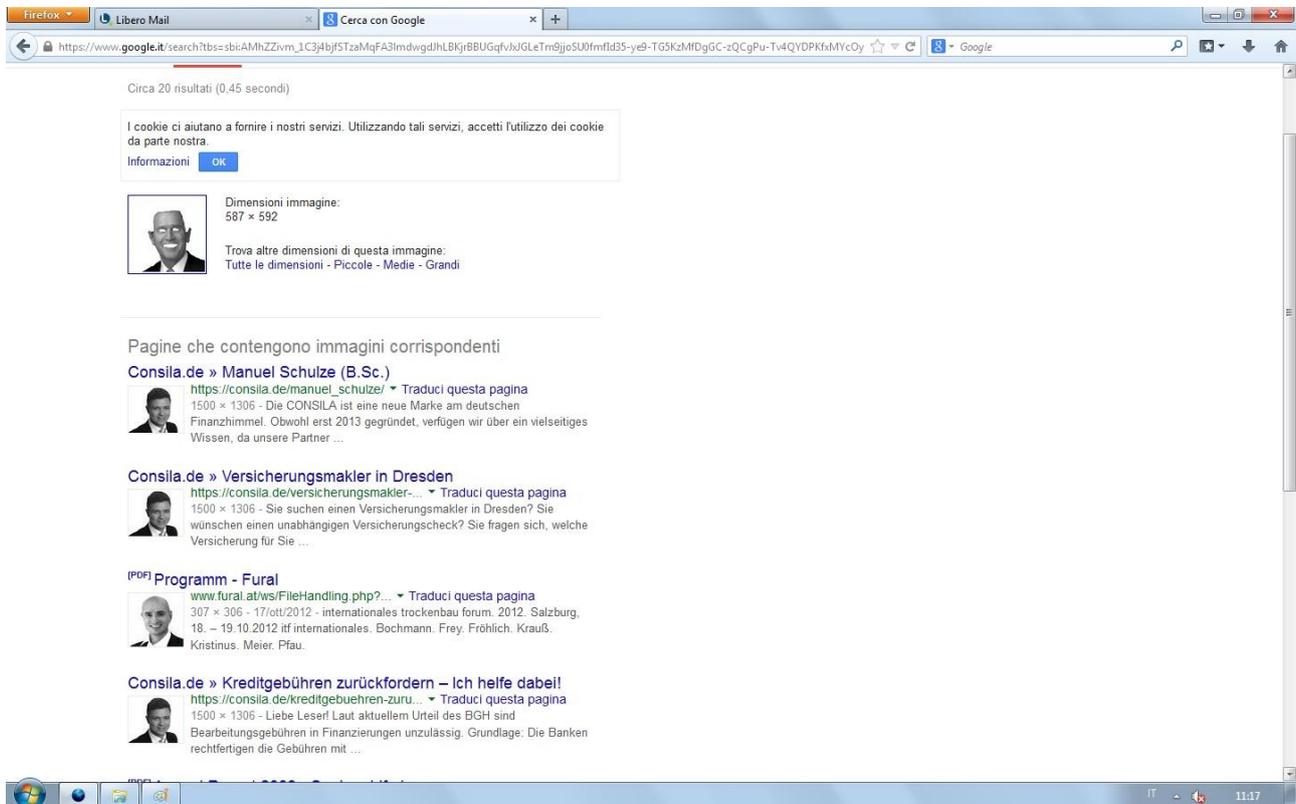


Fig.42 Risultati ottenuti da Google per l'immagine 6.

Immagine 7: Mettendo uno sfondo grigio, Google dà risultati coerenti e non coerenti. Si può ipotizzare che uno dei parametri che Google utilizza per classificare le immagini sia la media calcolata sui toni di grigio. L'immagine trovata è a colori.

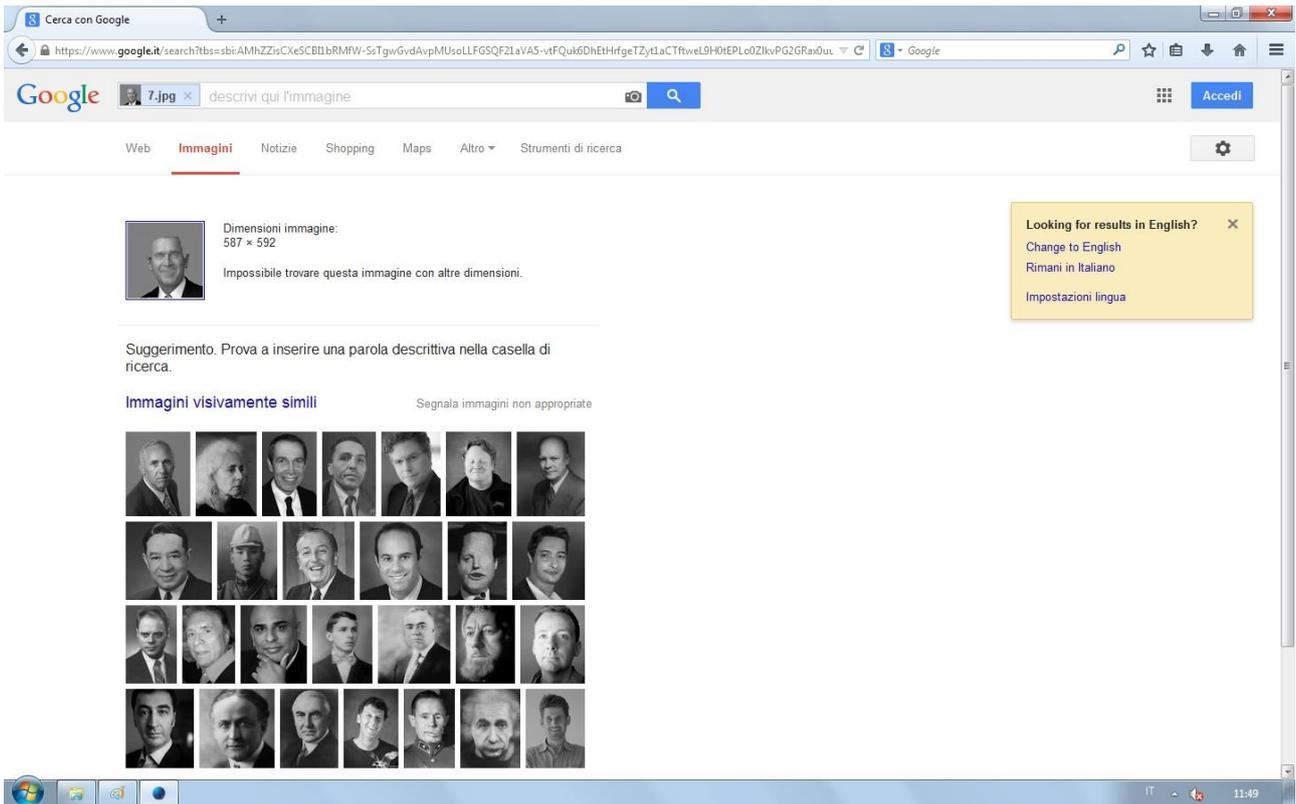


Fig.43 Risultato non coerente dato da Google per l'immagine 7.

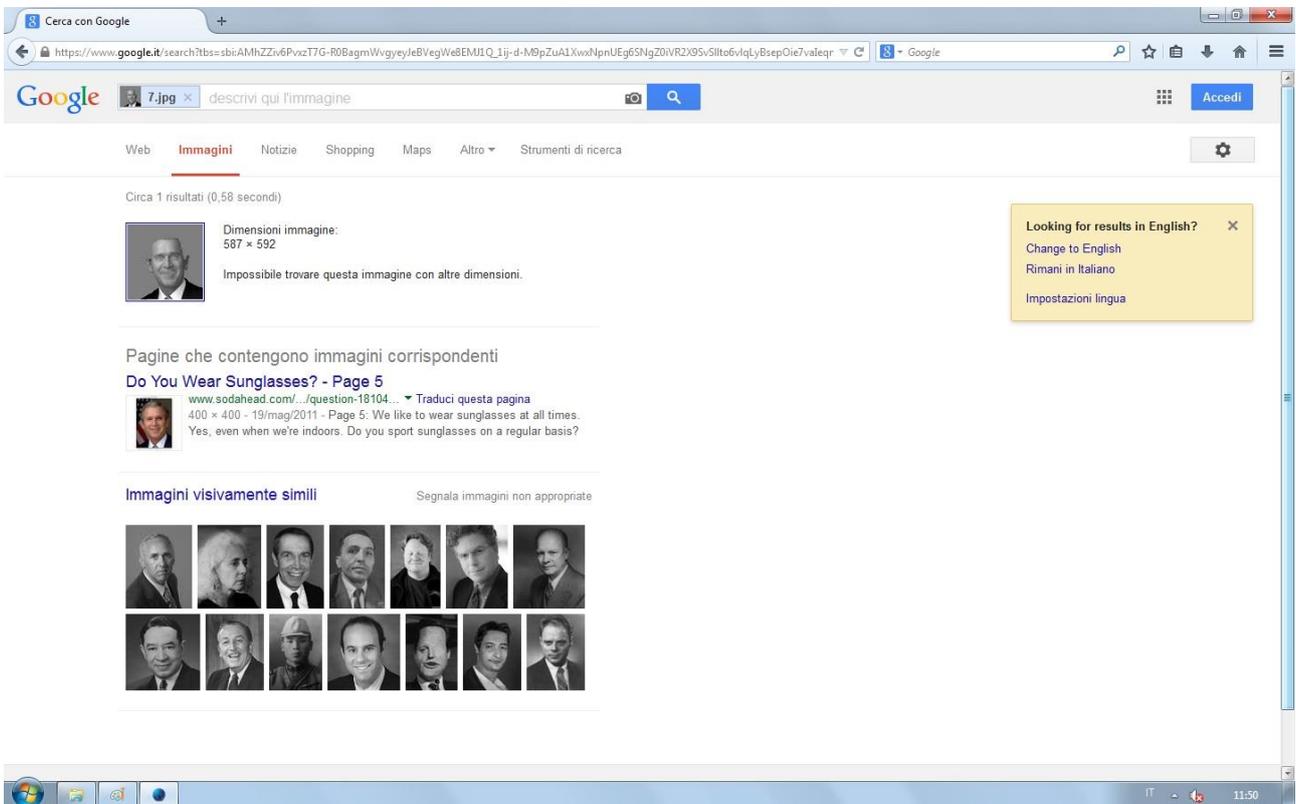


Fig.44 Risultato coerente dato da Google per l'immagine 7.

Immagine 8, 9, 10, 11, 12, 13: Google riesce a dare risultati coerenti e non coerenti per l'immagine 10, coerenti per le immagini 11 e 12, nessun risultati per la 8, 9 e 13. La coerenza dei risultati dati da Google è dovuta allo sfondo grigio.

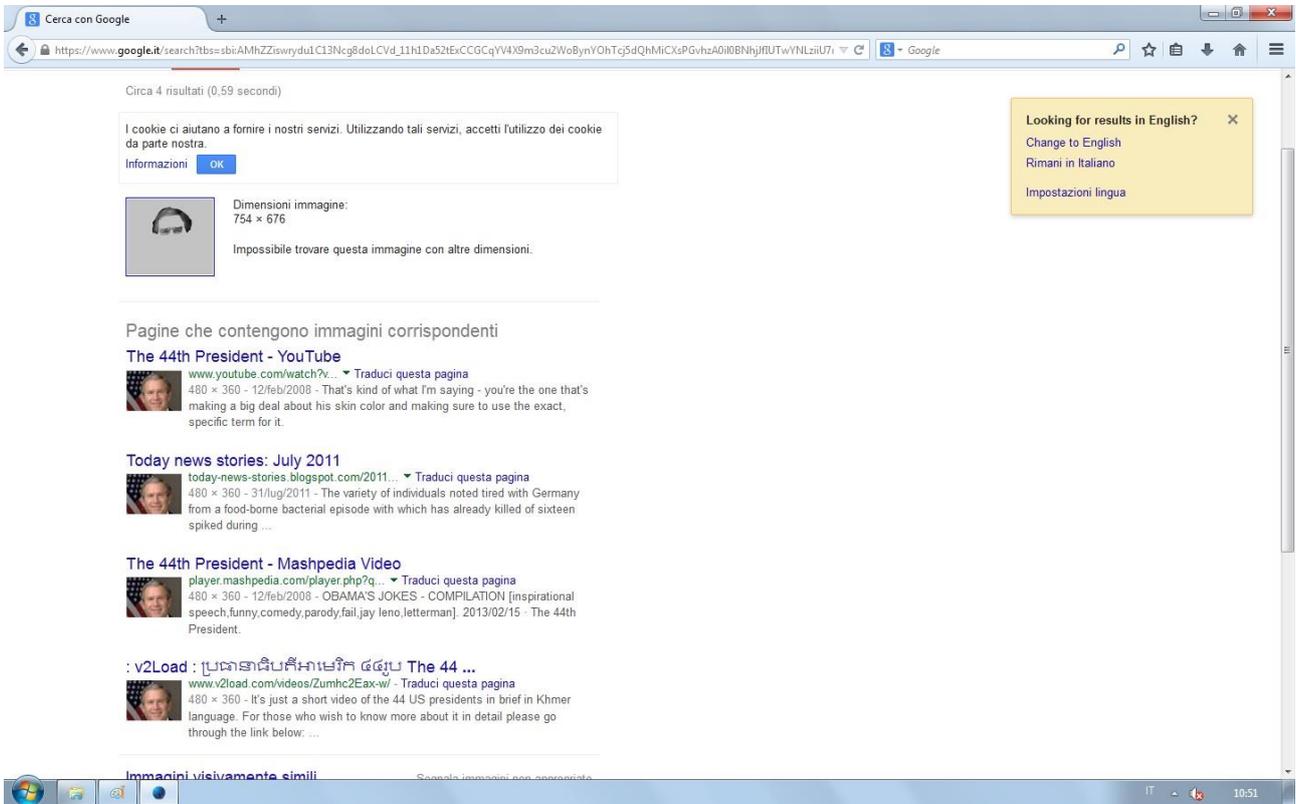


Fig.45 Risultato coerente fornito da Google per l'immagine 10.

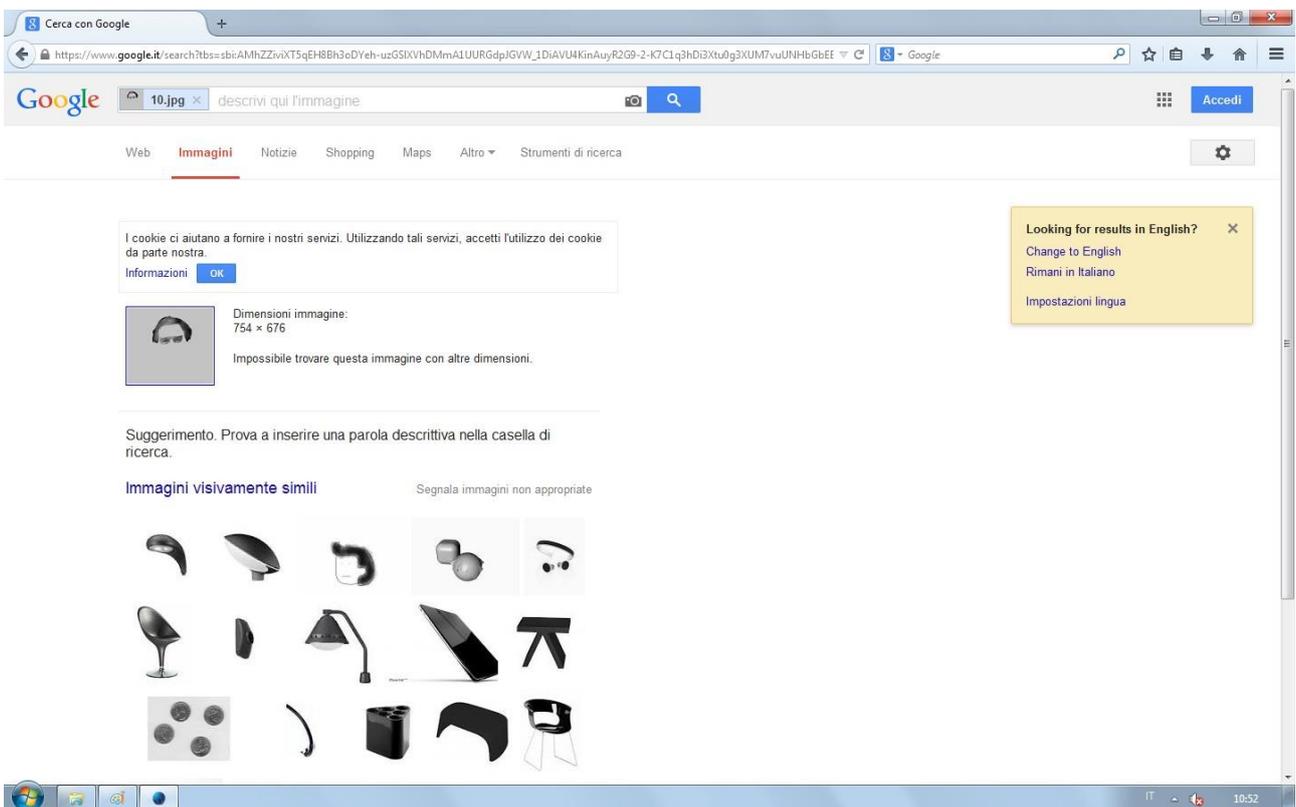


Fig.46 Risultato non coerente fornito da Google per l'immagine 10.

Immagini 14, 15, 16: Inclinando orizzontalmente l'immagine di 10°, 20° e 30° si vuol vedere se Google trova risultati con il volto inclinato o se ha in archivio un'immagine simile.

	Google
Imm.14	L'immagine così inclinata è in archivio.
Imm.15	Si hanno risultati coerenti e incoerenti
Imm.16	Non dà risultato

Fig.47 Tabella dei risultati ottenuti da Google per le immagini 14, 15, 16.

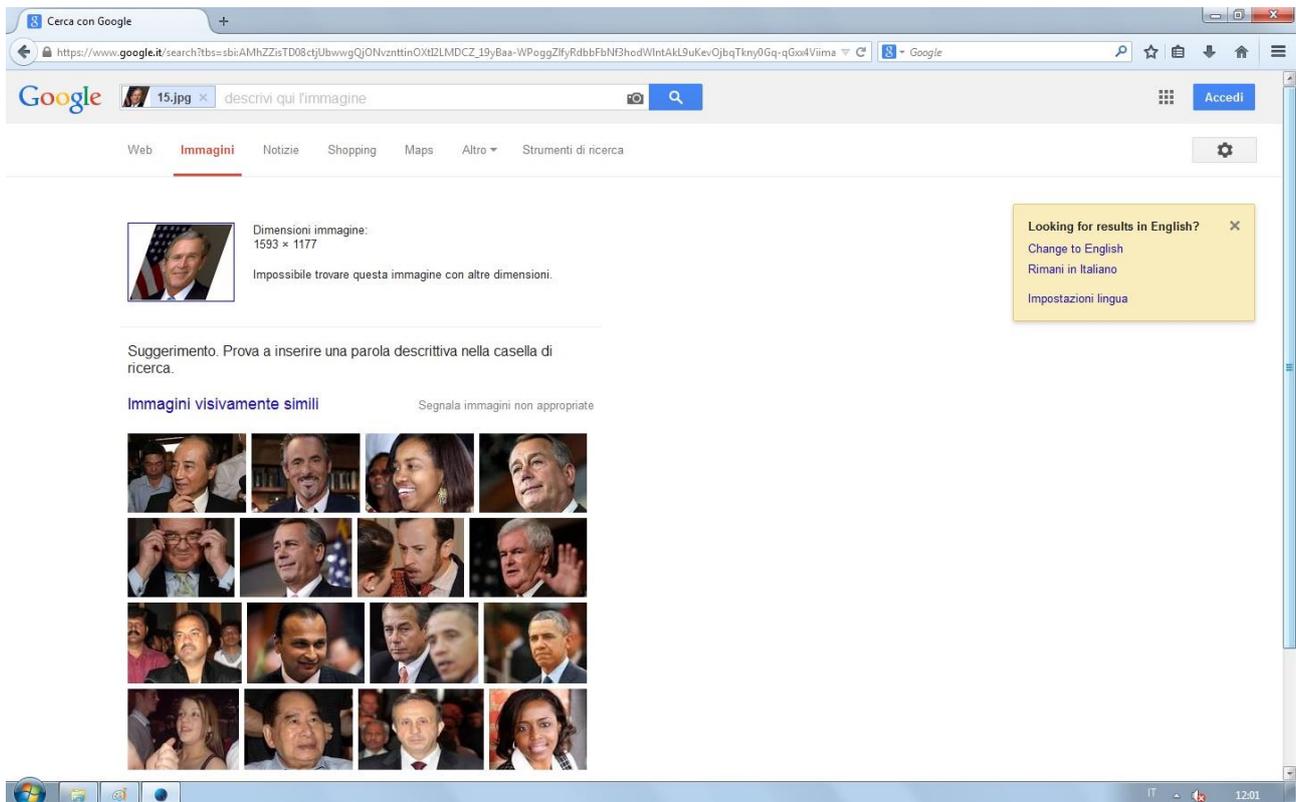


Fig.48 Risultato non coerente ottenuto da Google per l'immagine 15.

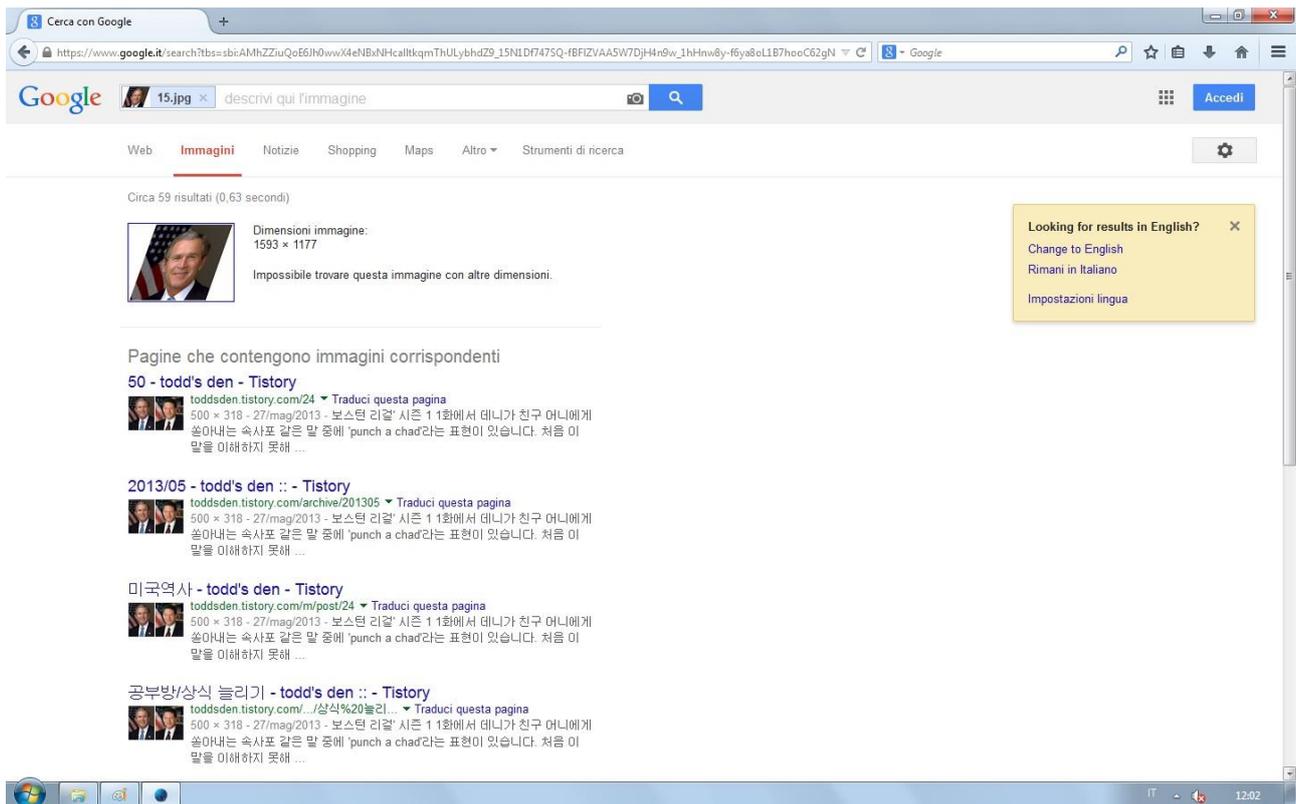


Fig.49 Risultato coerente ottenuto da Google per l'immagine 15.

Google riconosce l'operazione di inclinazione fatta nell'immagine 15 come un rimpicciolimento.

Immagini 17,18: rotando l'immagine 1 rispettivamente di 90° e 180° si vuol vedere se Google dà come risultati immagini rotate o non rotate. I risultati forniti dal motore di ricerca per l'immagine rotata di 90° non hanno nulla di particolare mentre sono significativi i risultati per l'immagine rotata di 180°. Google riconosce che il viso è rotato di 180° dando come risultati immagini in cui lo sfondo è sia rotato che non rotato.

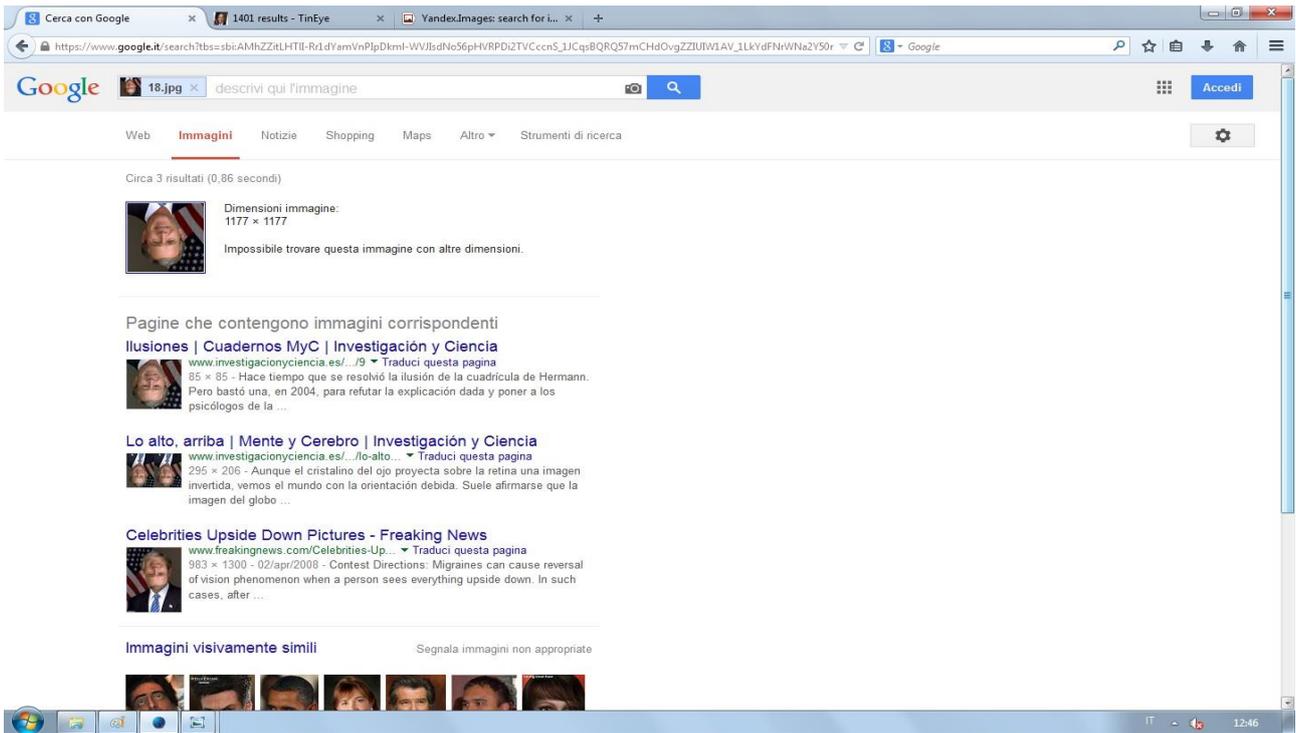


Fig.50 Risultato ottenuto da Google sottoponendo l'immagine 18.

Immagine 19: Google dà risultati non coerenti come nell'immagine 3; tuttavia i risultati sono differenti, questo per la presenza del colore che nell'immagine 3 è assente.

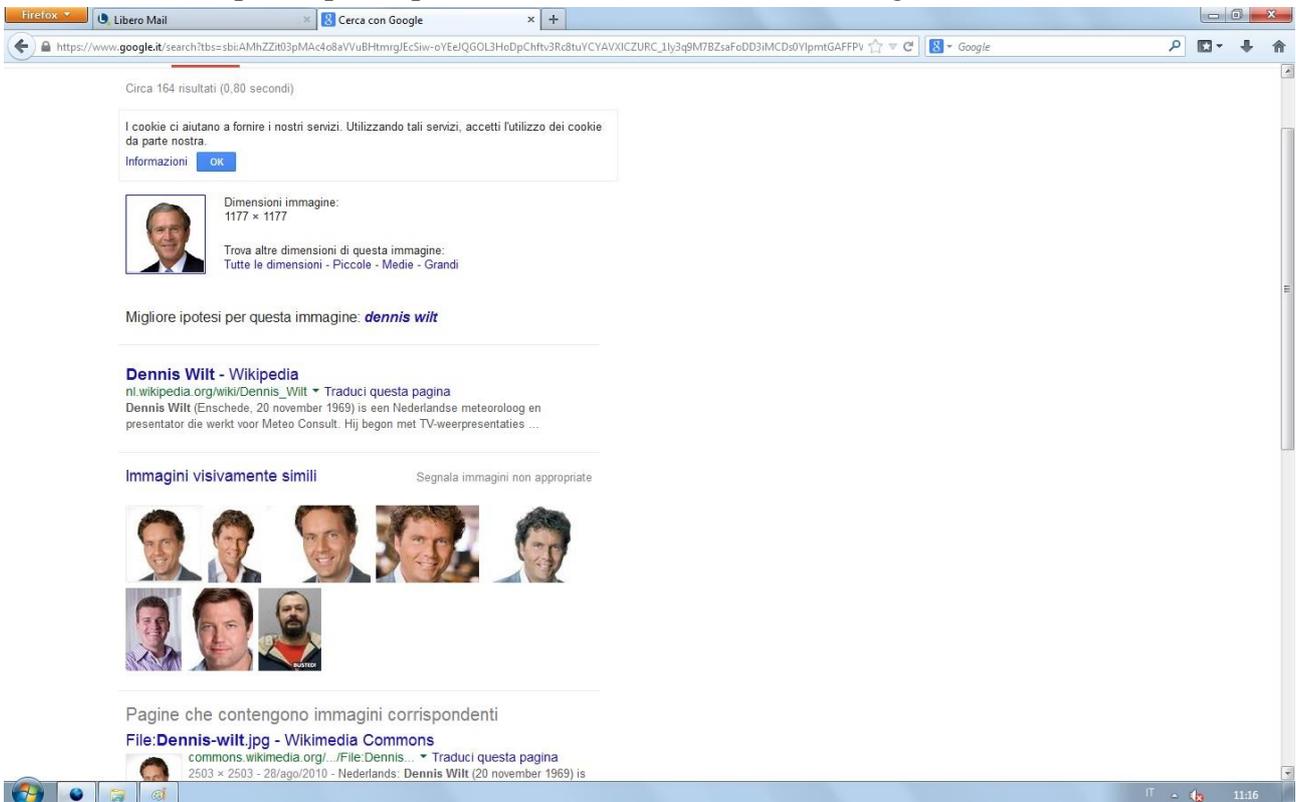


Fig.51 Risultato ottenuto da Google sottoponendo l'immagine 19.

Immagini 20,21: non si hanno risultati da Google.

Immagine 22, 23, 24: Google presenta risultati coerenti solo con l'immagine 24.



Fig.52 Risultato coerente ottenuto da Google sottoponendo l'immagine 24.

2.2 TinEye

2.2.1 Interfaccia utente

L'interfaccia che TinEye presenta all'utente è la più semplice: basta andare su www.tineye.com e trascinare l'immagine che si vuole cercare nella barra di ricerca. Il motore di ricerca indicherà il numero di risultati, il tempo impiegato, il numero di immagini su cui è stata fatta la ricerca e l'elenco dei siti dove possiamo trovare l'immagine da noi usata, modificata o originale.



Fig.53 Esempio di ricerca con TinEye.

A sinistra si notano diverse operazioni di filtraggio: per data, per dimensione, e per modifiche di immagine.

2.2.2 Risultati di TinEye per le immagini test in fig.7

Sottoponendo a TinEye le immagini test in fig.7 si ottengono i risultati riportati nella seguente tabella.

Mot.Ricerca Immagini	TinEye	
	Numero Risultati	Match
Imm.1	2464	Sì
Imm.2	2342	Sì
Imm.3	1079	Sì
Imm.4	83	Sì
Imm.5	33	Sì

Imm.6	26	Sì
Imm.7	36	Sì
Imm.8	0	No
Imm.9	0	No
Imm.10	0	No
Imm.11	0	No
Imm.12	0	No
Imm.13	0	No
Imm.14	1039	Sì
Imm.15	0	No
Imm.16	0	No
Imm.17	1	Sì
Imm.18	1	Sì
Imm.19	636	Sì
Imm.20	0	No
Imm.21	0	No
Imm.22	88	Sì
Imm.23	36	Sì
Imm.24	26	Sì

Fig.54 Tabella risultati per le immagini test in fig.7

Per TinEye il numero di risultati rappresenta il numero di immagini recuperate nel suo database.

Immagine 1: TinEye dà risultati coerenti. In questo caso l'immagine originale a colori dà il massimo dei risultati.

Immagine 2: la perdita dell'informazione cromatica influenza poco TinEye. Il numero di risultati varia molto poco e le immagini trovate sono tutte a colori e coerenti. Si può ipotizzare che l'istogramma di colore ha poca influenza nell'algoritmo di classificazione delle immagini usato da TinEye

Immagine 3: la perdita dello sfondo dimezza i risultati, che ancora rimangono coerenti.

Immagine 4, 5, 6: la perdita degli occhi è quella che riduce maggiormente i risultati, rispetto a bocca e capelli; i risultati rimangono coerenti.

Immagine 7: Mettendo uno sfondo grigio si vede che TinEye trova più immagini. Si può ipotizzare che uno dei parametri che TinEye utilizza per classificare le immagini sia la media calcolata sui toni di grigio. L'immagine trovata è a colori.

Immagine 8, 9, 10, 11, 12, 13: TinEye non dà risultati: la perdita di informazione è troppo elevata.

Immagini 14, 15, 16: Inclinando orizzontalmente l'immagine di 10°, 20° e 30° si vuol vedere se i motori di ricerca trovano risultati con il volto inclinato o se hanno in archivio un'immagine simile. Per l'immagine 14 TinEye dà risultati coerenti, non inclinate e sempre a colori. Mentre per l'immagine 15 e 16 non dà alcun risultato.

	TinEye
Imm.14	Immagini non inclinate.
Imm.15	Nessun risultato
Imm.16	Nessun risultato

Fig.55 Tabella dei risultati ottenuti da TinEye per le immagini 14, 15, 16.

Immagini 17,18: rotando l'immagine 1 rispettivamente di 90° e 180° si vuol vedere se TinEye dà come risultati immagini rotate o non rotate. I risultati forniti da TinEye per l'immagine rotata di 90° non hanno nulla di particolare mentre sono significativi i risultati

per l'immagine rotata di 180°. TinEye dà come unico risultato un'immagine in cui è solo lo sfondo ad essere rotato invece il viso è in posizione regolare.

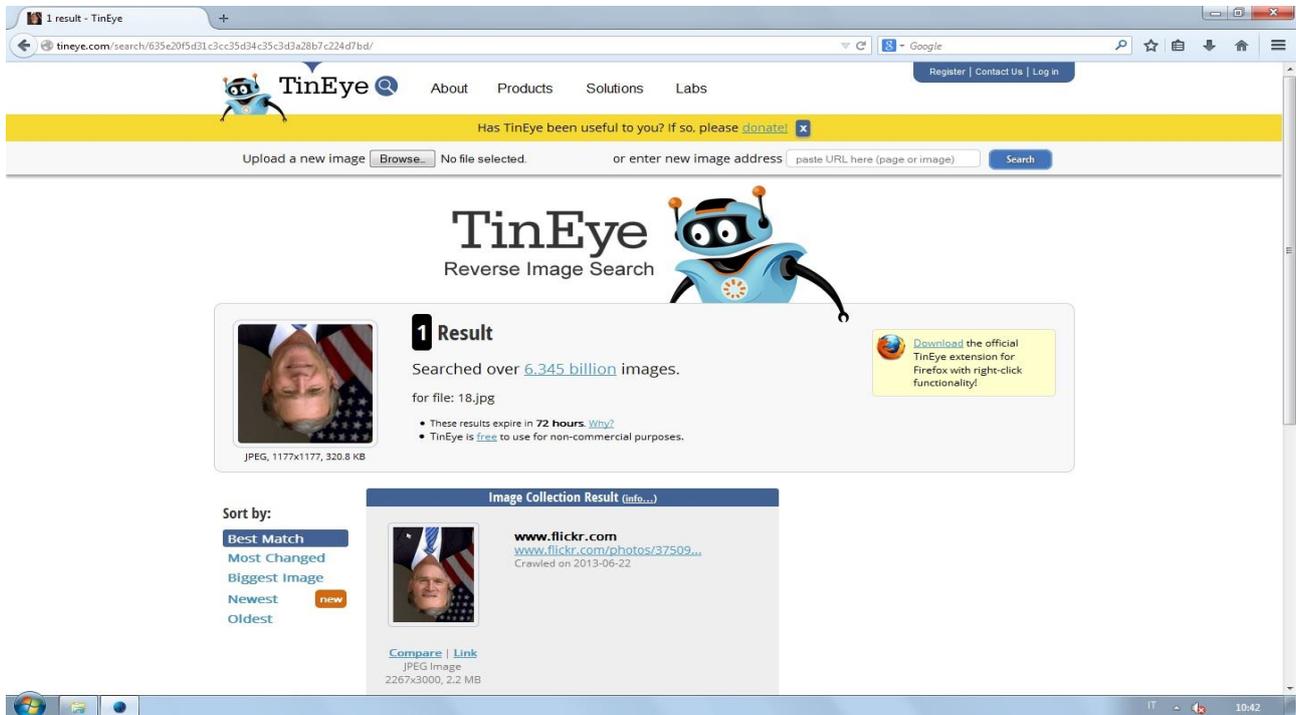


Fig.56 Risultato ottenuto da TinEye sottoponendo l'immagine 18.

Immagine 19: TinEye dà risultati coerenti.

Immagini 20,21: non si hanno risultati da TinEye.

Immagini 22, 23, 24: TinEye presenta risultati analoghi a quelli dati per le immagini 4, 5, 6.

2.3 Yandex

2.3.1 Interfaccia utente

Il motore di ricerca Yandex (www.yandex.com) presenta subito una barra di ricerca testuale con diverse opzioni: basta cliccare su immagini e trascinare l'immagine che vogliamo cercare nel web sulla barra di ricerca in alto. Yandex è il motore di ricerca più povero a livello di contenuti: ci indica al fianco dell'immagine test le dimensioni in cui possiamo trovarla e i siti in cui compare ma non fornisce il numero di risultati trovati, quindi si potrà solo dire se i risultati saranno coerenti o non coerenti con la ricerca fatta.

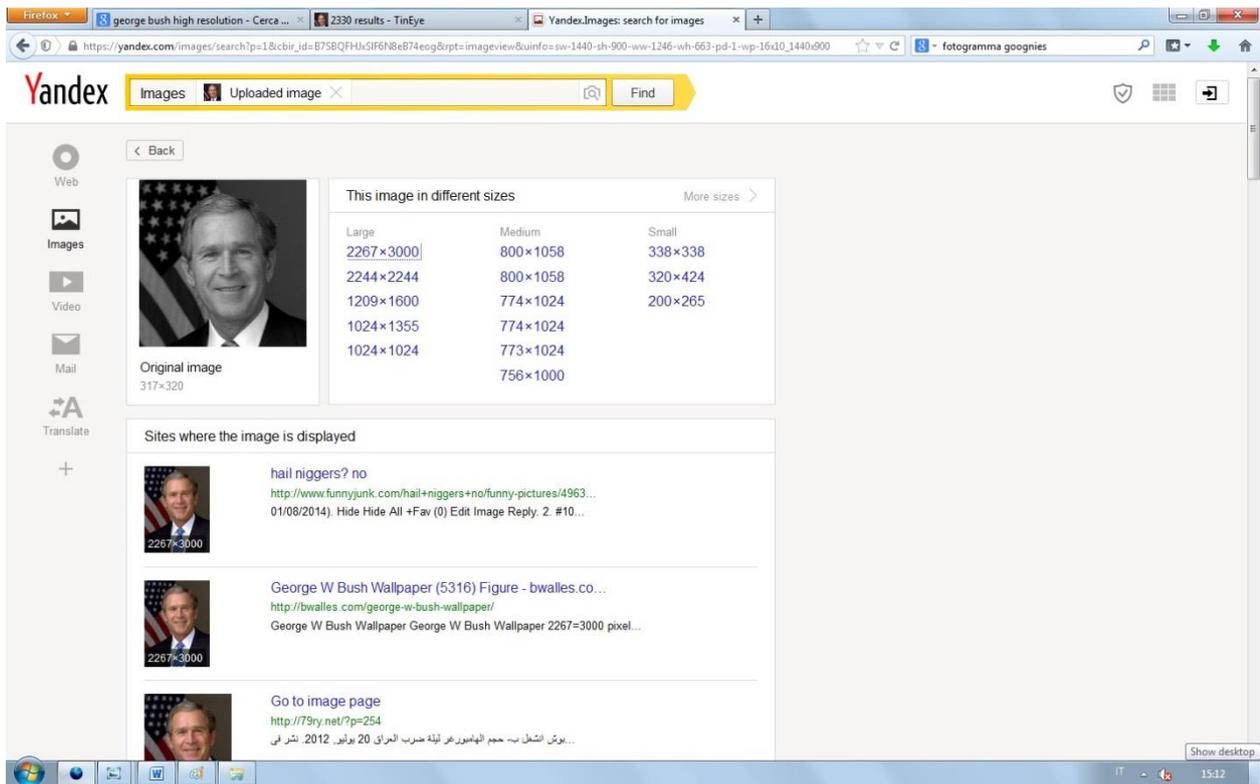


Fig.57 Esempio di ricerca con Yandex.

2.3.2 Risultati di Yandex per le immagini test in fig.7

Sottoponendo a Yandex le immagini test in fig.7 si ottengono i risultati riportati nella seguente tabella.

Mot.Ricerca	Yandex	
Immagini	Match	
	Imm.1	Sì
	Imm.2	Sì
	Imm.3	Sì
	Imm.4	Sì
	Imm.5	Sì
	Imm.6	Sì

Imm.7	Sì
Imm.8	No
Imm.9	No
Imm.10	No
Imm.11	No
Imm.12	No
Imm.13	No
Imm.14	Sì
Imm.15	Sì
Imm.16	No
Imm.17	Sì
Imm.18	Sì
Imm.19	Sì
Imm.20	Dà risultati ma non pertinenti
Imm.21	Dà risultati ma non pertinenti
Imm.22	Sì
Imm.23	Sì
Imm.24	Sì

Fig.58 Tabella risultati per le immagini test in fig.7

Nella tabella è stata elencata solo la coerenza dei risultati in quanto Yandex non mette a disposizione dell'utente il numero dei risultati ottenuti con la ricerca. Le ricerche sono state fatte utilizzando finestre anonime in quanto Yandex utilizza i cookie.

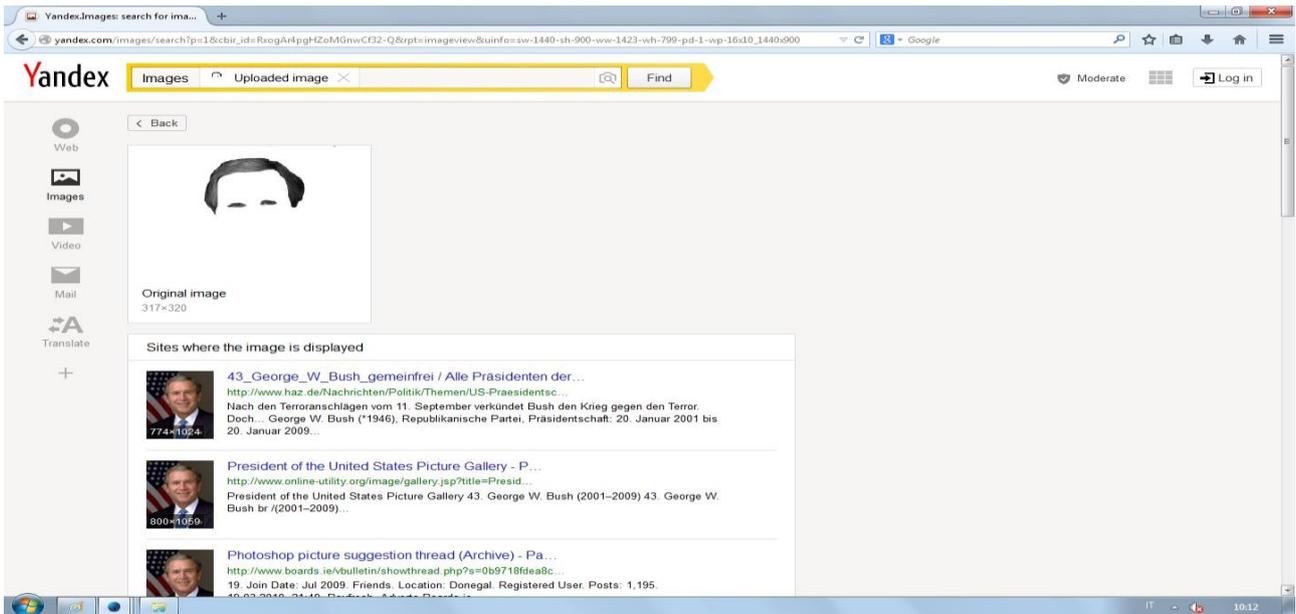


Fig.59 Risultato presentato da Yandex avendo fatto ricerche precedenti con l'immagine di George Bush.

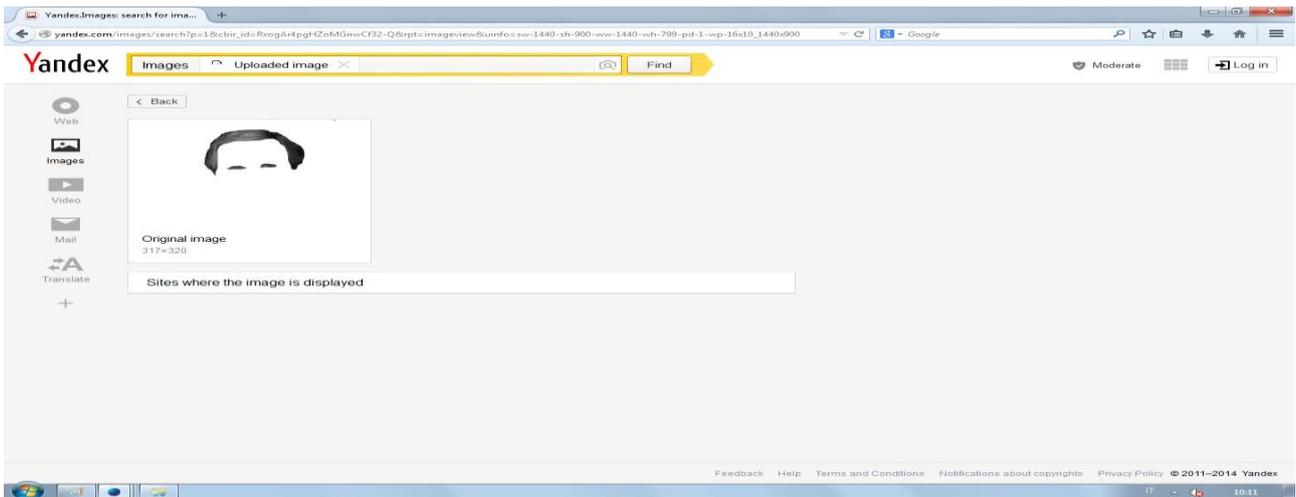


Fig.60 Risultato presentato da Yandex utilizzando la finestra anonima.

L'algoritmo di recupero delle immagini di Yandex sembra essere molto simile a quello di TinEye. Le differenze con TinEye sono i risultati per le immagini 15,18,20,21.

Immagini 14, 15, 16: Inclinando orizzontalmente l'immagine di 10°, 20° e 30° si vuol vedere se i motori di ricerca trovano risultati con il volto inclinato o se hanno in archivio un'immagine simile.

Yandex	
Imm.14	Immagini non inclinate.

Imm.15	Immagini non inclinate.
Imm.16	Nessun risultato

Fig.61 Tabella dei risultati ottenuti da Yandex per le immagini 14, 15, 16.

Yandex dà risultati fino ad una inclinazione di 20°.

Immagini 17,18: rotando l'immagine 1 rispettivamente di 90° e 180° si vuol vedere se i motori di ricerca danno come risultati immagini rotate o non rotate. I risultati forniti dai tre motori di ricerca per l'immagine rotata di 90° non hanno nulla di particolare mentre sono significativi i risultati per l'immagine rotata di 180°. Yandex riconosce che il viso è rotato di 180° dando come risultati immagini in cui lo sfondo è sia rotato che non rotato.

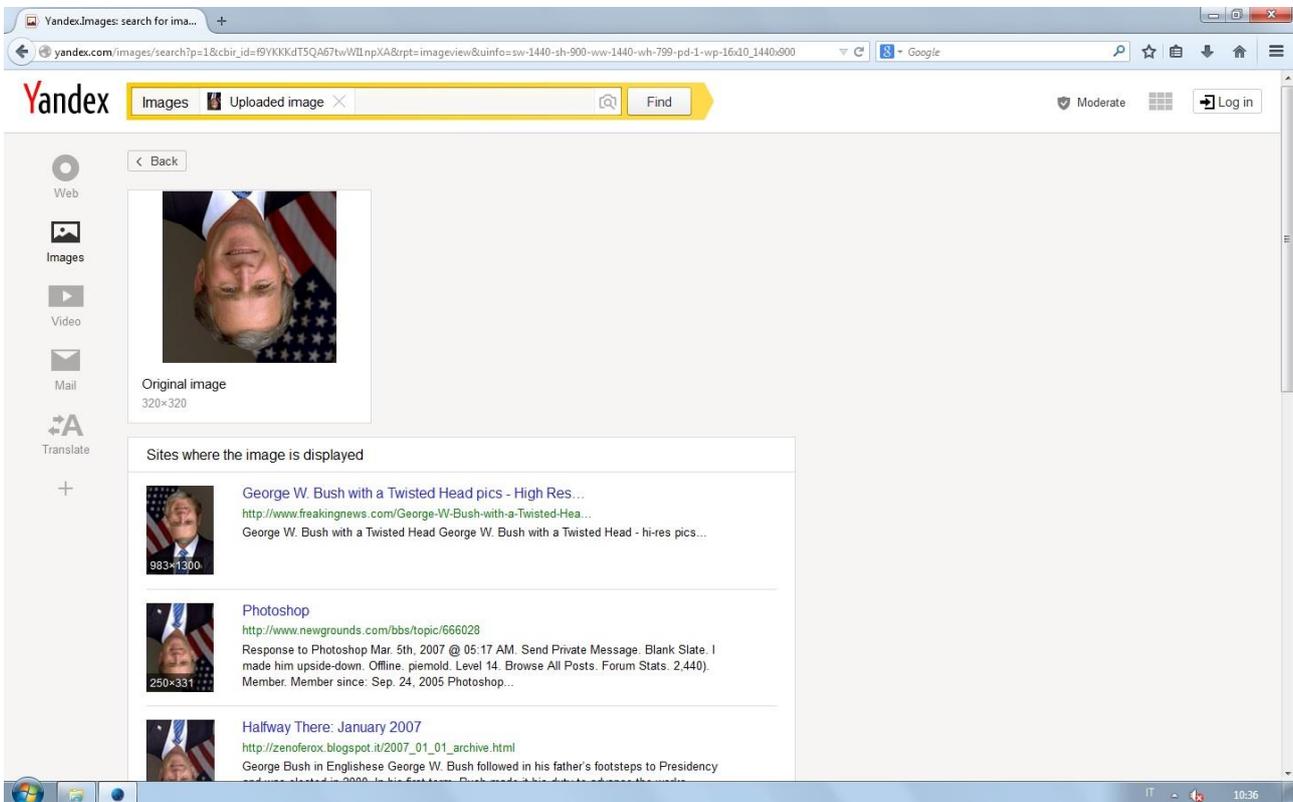


Fig.62 Risultato ottenuto da Yandex sottoponendo l'immagine 18.

Immagini 20,21: Yandex dà risultati non coerenti. Si può notare che Yandex non utilizzi l'istogramma di colore come informazione primaria e si può ipotizzare che utilizzi un algoritmo che calcola il rettangolo di base.

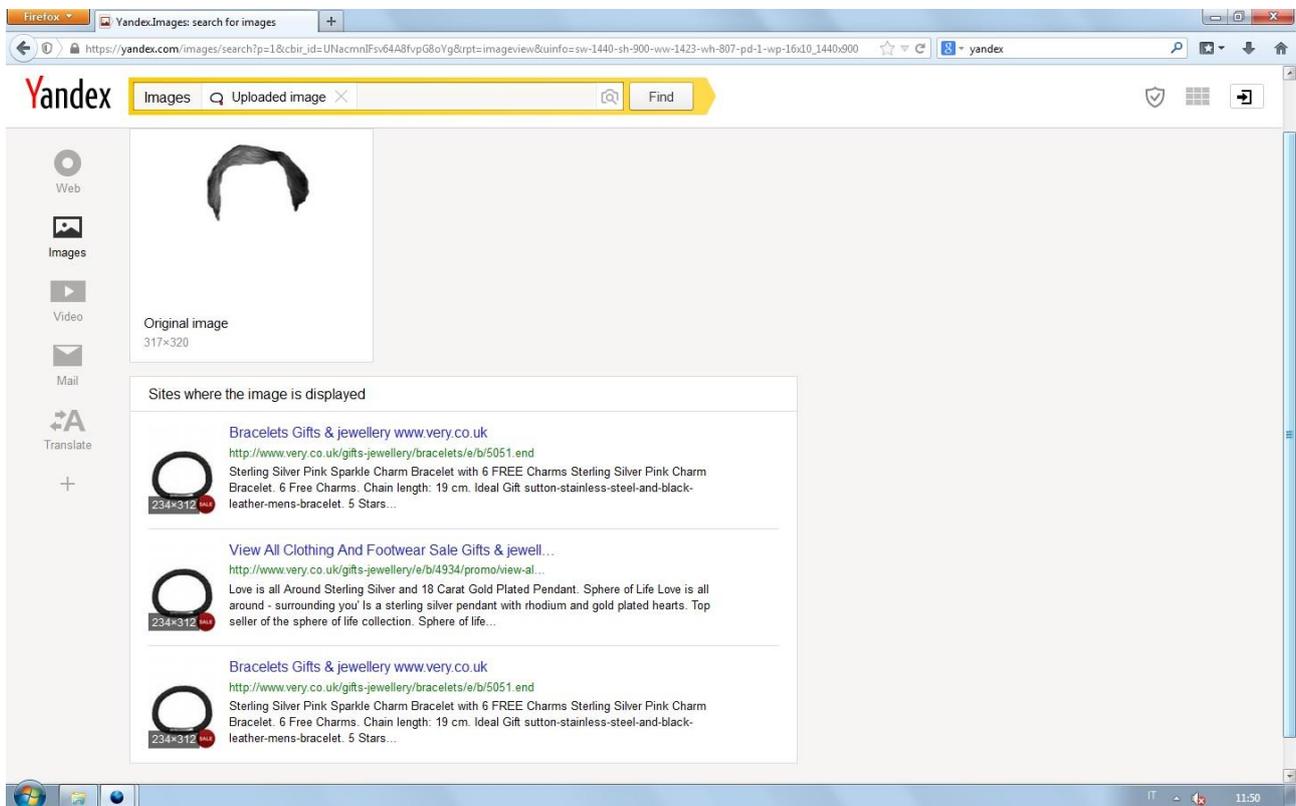


Fig.63 Risultato non coerente ottenuto da Yandex sottoponendo l'immagine 20.

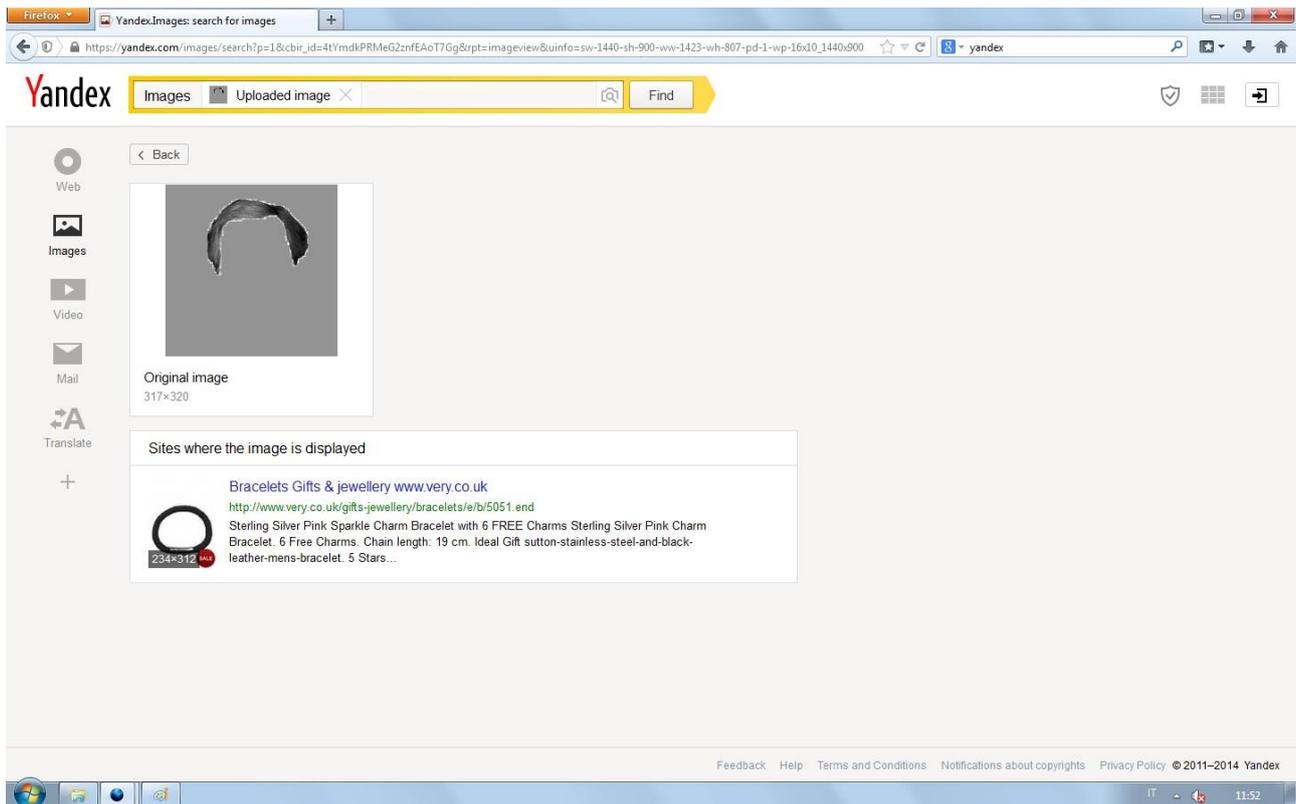


Fig.64 Risultato non coerente ottenuto da Yandex sottoponendo l'immagine 21.

Conclusione

Dall'osservazione dei tre motori di ricerca possiamo concludere che Google è il motore di ricerca sicuramente più completo. Fornisce all'utente un servizio che nessun altro motore di ricerca offre, ovvero il poter cercare immagini correlate a quella che gli si è sottoposta. Tuttavia Google risulta meno performante rispetto agli altri motori di ricerca osservati a causa della presentazione random dei risultati, e per il fatto che l'algoritmo di ricerca per le immagini visivamente simili influenza l'algoritmo di ricerca dell'immagine nel web. TinEye e Yandex danno risultati simili, ma a mio avviso il migliore è TinEye perché ha una migliore interfaccia utente rispetto a Yandex e per il fatto che Yandex ha dato risultati non coerenti per le immagini 20,21.

Bibliografia

- [1] M. Stricker, and M. Orengo, "Similarity of color images," SPIE Storage and Retrieval for Image and Video Databases III, vol. 2185, pp.381-392.
- [2] H. Tamura, S. Mori, and T. Yamawaki, "Texture features corresponding to visual perception," IEEE Trans. On Systems, Man, and Cybernetics, vol. Smc-8, No. 6.
- [3] J. M. Francos. "Orthogonal decompositions of 2D random fields and their applications in 2Dspectral estimation," N. K. Bose and C. R. Rao, editors, Signal Processing and its Application, pp.20-227. North Holland.
- [4] J. Mao, and A. K. Jain, "Texture classification and segmentation using multiresolution simultaneous autoregressive models," Pattern Recognition, Vol. 25, No. 2, pp. 173-188.
- [5] J. G. Daugman, "Complete discrete 2D Gabor transforms by neural networks for image analysis and compression," IEEE Trans. ASSP, vol. 36 pp. 1169-1179.
- [6] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," IEEE Trans. on Information Theory, Vol. 36, pp 961-1005.
- [7] J. E. Gary, and R. Mehrotra, "Shape similarity-based retrieval in image database systems," Proc. Of SPIE, Image Storage and Retrieval Systems, vol. 1662 pp. 2-8.
- [8] S. K. Chang, Q. Y. Shi, and C. Y. Yan, "Iconic indexing by 2-D strings," IEEE Trans. on Pattern Anal. Machine Intell., Vol.9, No.3, pp. 413-428, May 1987.

