

Speeded-Up Robust Features: Un'Analisi dell'Algoritmo SURF

Leonardo Fusar Bassini

Dipartimento di Matematica, Università Statale di Milano, Italia

Giugno 2025

Sommario

- 1 Introduzione: La Sfida della Visione
- 2 Sfide: La Ricerca dell'Invarianza
- 3 SIFT: Un Punto di Riferimento
- 4 SURF: Fase 1 - Immagini Integrali
- 5 SURF: Fase 2 - Rilevamento Punti di Interesse
- 6 SURF: Fasi 3 e 4 - Orientamento e Descrittore
- 7 Implementazione e Verifica Sperimentale
- 8 Conclusione

Abstract: Punti Chiave

- Sfide legate a norma e proprietà topologiche delle immagini.
- Innovazioni computazionali dell'algoritmo SURF:
 - Immagini integrali per elaborazione accelerata.
 - Approssimazione efficiente della matrice Hessiana per rilevamento punti di interesse.
 - Risposte di wavelet di Haar per descrittori compatti e invarianti all'orientamento.
- Verifica sperimentale di come vengono "imitati" i meccanismi d'attenzione
- Spiega come SURF imita l'attenzione umana per la corrispondenza e la comprensione della scena.

Introduzione: La Sfida della Visione

- **Fragilità Topologica delle Immagini:**

- Piccole perturbazioni (rumore, luminosità, rotazioni) alterano drasticamente i pixel.
- Necessità di strutture invarianti, confronti pixel-wise e squared-error inaffidabili.

- **Oltre i Pixel: Comprensione Semantica**

- La percezione umana è contestuale.
- La visione artificiale deve affrontare una comprensione di livello superiore.
- Descrittori locali come SURF ricavano blocchi costitutivi più robusti a livello semantico.

- **Gli approcci tradizionali spesso falliscono;** SIFT è spesso troppo lento per applicazioni in tempo reale o grandi database.

"Fragilità topologica"

Published as a conference paper at ICLR 2015

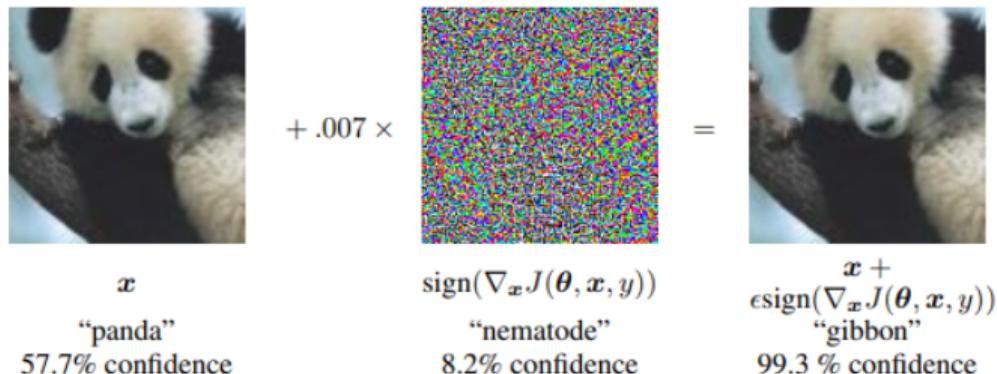


Figure 1: A demonstration of fast adversarial example generation applied to GoogLeNet (Szegedy et al., 2014a) on ImageNet. By adding an imperceptibly small vector whose elements are equal to the sign of the elements of the gradient of the cost function with respect to the input, we can change GoogLeNet's classification of the image. Here our ϵ of .007 corresponds to the magnitude of the smallest bit of an 8 bit image encoding after GoogLeNet's conversion to real numbers.

EXPLAINING AND HARNESSING ADVERSARIAL EXAMPLES, Ian J. Goodfellow, Jonathon Shlens & Christian Szegedy

Sfide: La Ricerca dell'Invarianza

- **Individuazione Punti Salienti: "Attenzione" e "Focus"**

- Trovare punti/regioni informative ripetibili.
- Guidare l'attenzione del computer.

- **Raggiungere una Robusta Invarianza**

- **Scala:** Indipendente dalle dimensioni.
- **Rotazione:** Indipendente dall'orientamento.
- **Illuminazione:** Robusto a luminosità, contrasto, ombre. **Punto di Vista:** Robustezza a leggere trasformazioni affini.
- **Rumore:** Stabile anche con rumore.

- **Descrittori Compatti e Discriminanti**

- "Impronta digitale" unica e coerente.
- Efficienza computazionale.

Immagini Contestuali



Figura: Immagine di una macchina

Immagini Contestuali



Figura: Pubblicità di una macchina

Immagini Contestuali

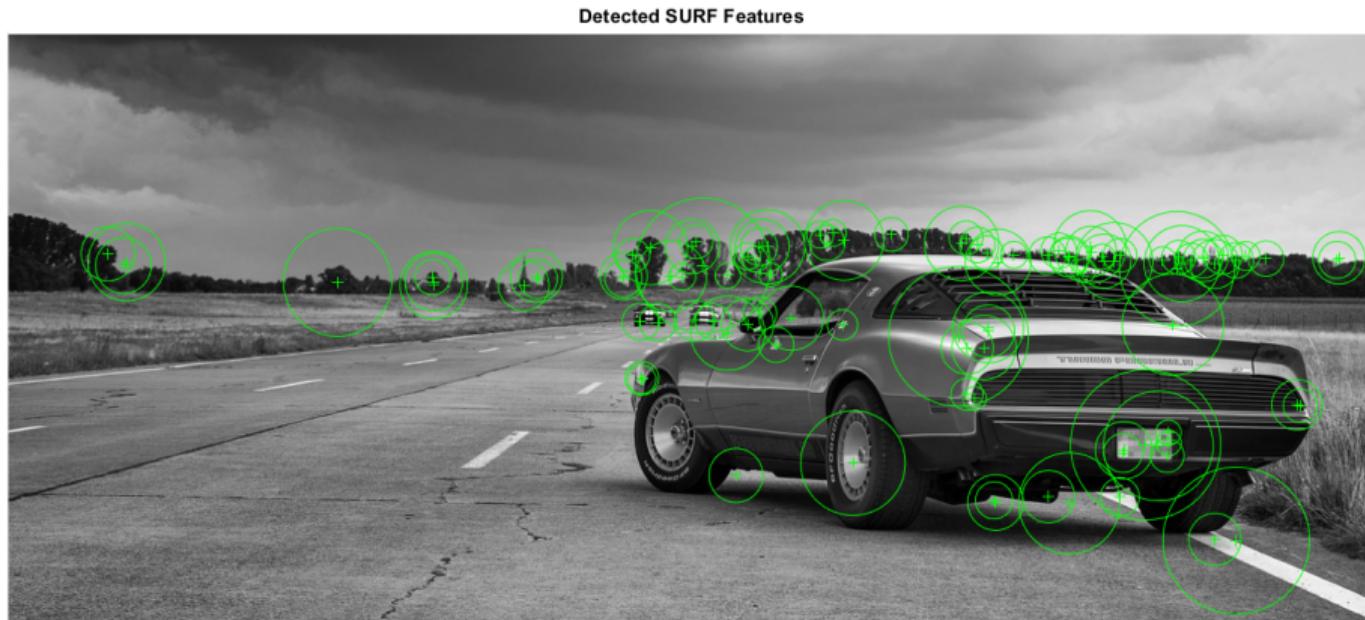


Figura: Immagine di una macchina

Immagini Contestuali



Figura: Pubblicità di una macchina

SIFT: Un Punto di Riferimento

- SIFT ha rivoluzionato il campo, introducendo uno strumento resistente a piccole variazioni "naturali".
- **Rilevamento Punti Chiave nello Spazio di Scala:**
 - Trova punti chiave stabili a diverse scale.
 - Costruisce uno **spazio di scala** con smoothing gaussiano.
 - Identifica gli estremi nelle immagini **Difference-of-Gaussians (DoG)**.
- **Output Punti Chiave:**
 - Posizione (x, y), Scala (σ), e Orientamento (θ).
 - L'orientamento dominante garantisce **invarianza di rotazione**.
 - Descrittore: Vettore a 128 dimensioni (gradienti locali), robusto.
- SIFT ha creato un robusto **feature extractor**, ma è computazionalmente intenso.

SURF: Fase 1 - Immagini Integrali

- L'immagine integrale è fondamentale per la velocità di SURF (preparazione $O(n^2)$, calcolo $O(1)$ della somma rettangolare).
- **Come Funziona:**
 - Per $I(x, y)$, crea $II(x, y)$ memorizzando la somma dei pixel $I(i, j)$ per $i \leq x$ e $j \leq y$.
 - **Formula per $II(x, y)$**

$$II(x, y) = I(x, y) + II(x - 1, y) + II(x, y - 1) - II(x - 1, y - 1)$$

- **Somma di un rettangolo da (x_1, y_1) a (x_2, y_2) :**

$$\text{Somma} = II(x_2 + 1, y_2 + 1) - II(x_1, y_2 + 1) - II(x_2 + 1, y_1) + II(x_1, y_1)$$

Visualizzazione Immagine Integrale

1	2	3	4
5	6	7	8
9	1	2	3
4	5	6	7

Griglia Originale (I)



1	3	6	10
6	14	24	36
15	24	36	51
19	33	51	73

Immagine Integrale (II)

SURF: Fase 2 - Rilevamento Punti di Interesse (Massimi della Hessiana)

- Rileva punti "a forma di blob" approssimando il determinante della matrice Hessiana con filtri box.

- **Concetti Chiave:**

- **Derivate Seconde Approssimate (Filtri Box):**

- L_{xx}, L_{yy}, L_{xy} approssimati usando filtri box.
 - Risposte (D_{xx}, D_{yy}, D_{xy}) calcolate in $O(1)$ tramite immagini integrali.

- **Costruzione Spazio di Scala (Senza Ricampionamento):**

- Applica filtri box di *dimensioni crescenti* alla stessa immagine integrale.

- **Determinante dell'Approssimazione della Hessiana (0.9 fattore empirico):**

$$\det(\mathbf{H}_{\text{approx}}) \sim D_{xx}D_{yy} - (0.9D_{xy})^2$$

- **Soppressione Non-Massimale in 3D:**

- Candidati sono massimi locali in un vicinato $3 \times 3 \times 3$.
 - Interpolazione quadratica 3D per precisione sub-pixel/sub-scala.

SURF: Fasi 3 e 4 - Orientamento e Descrittore

- **Fase 3: Assegnazione dell'Orientamento (Invarianza alla Rotazione)**

- Definire una regione circolare attorno al punto chiave.
- Calcolare le risposte alle wavelet di Haar ($\sim dx, dy$).
- Sommare $\sim dx$ e $\sim dy$ all'interno di finestre angolari.
- L'orientamento dominante è il settore con la somma vettoriale più grande.

- **Fase 4: Formazione del Descrittore (64-Dimensionale)**

- Estrarre e ruotare una patch quadrata attorno al punto chiave.
- Dividere la patch ruotata in una griglia 4×4 .
- Per ogni sotto-regione, accumulare quattro somme: $\sum dx, \sum |dx|, \sum dy, \sum |dy|$.
- Concatenare i 16 vettori per il descrittore SURF finale a 64 dimensioni.
- Normalizzare il descrittore per robustezza all'illuminazione.

Immagini Richieste: Barca

Base Image



Detected SURF Features



Immagini Richieste: Elefanti

Base Image



Detected SURF Features



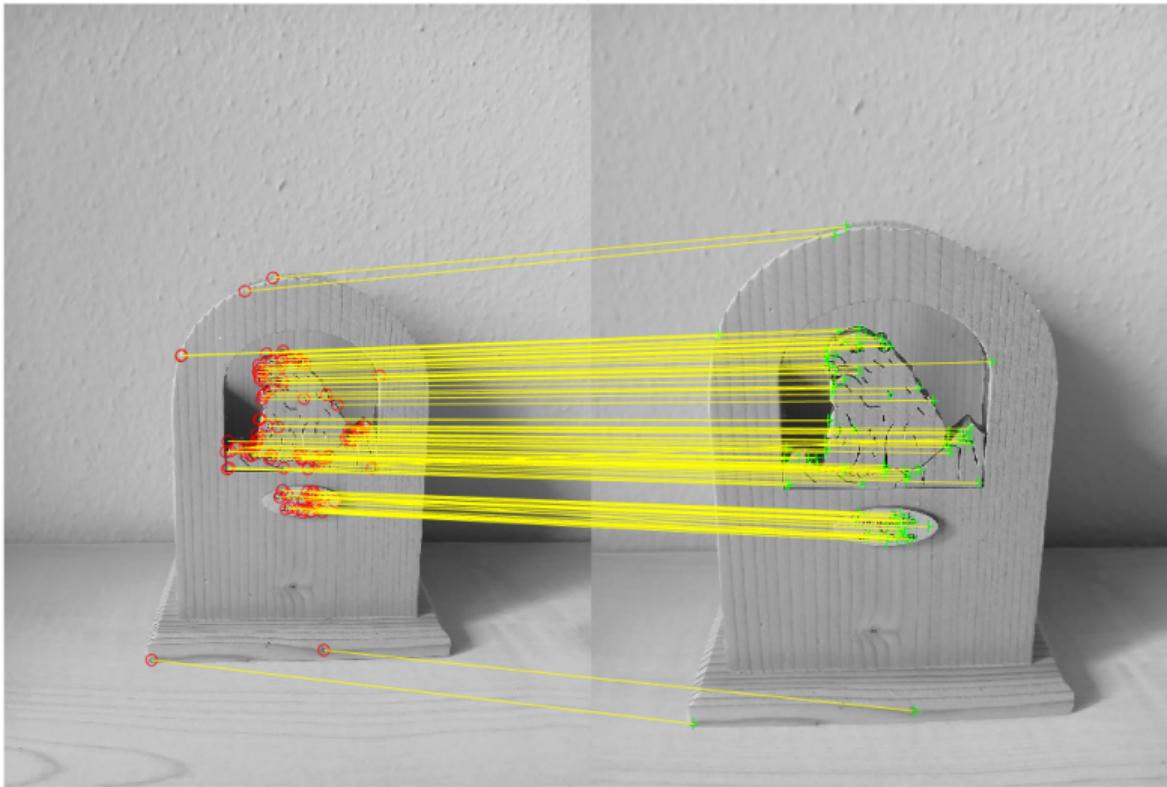
Implementazione e Verifica Sperimentale

- **Strategia di Implementazione:**

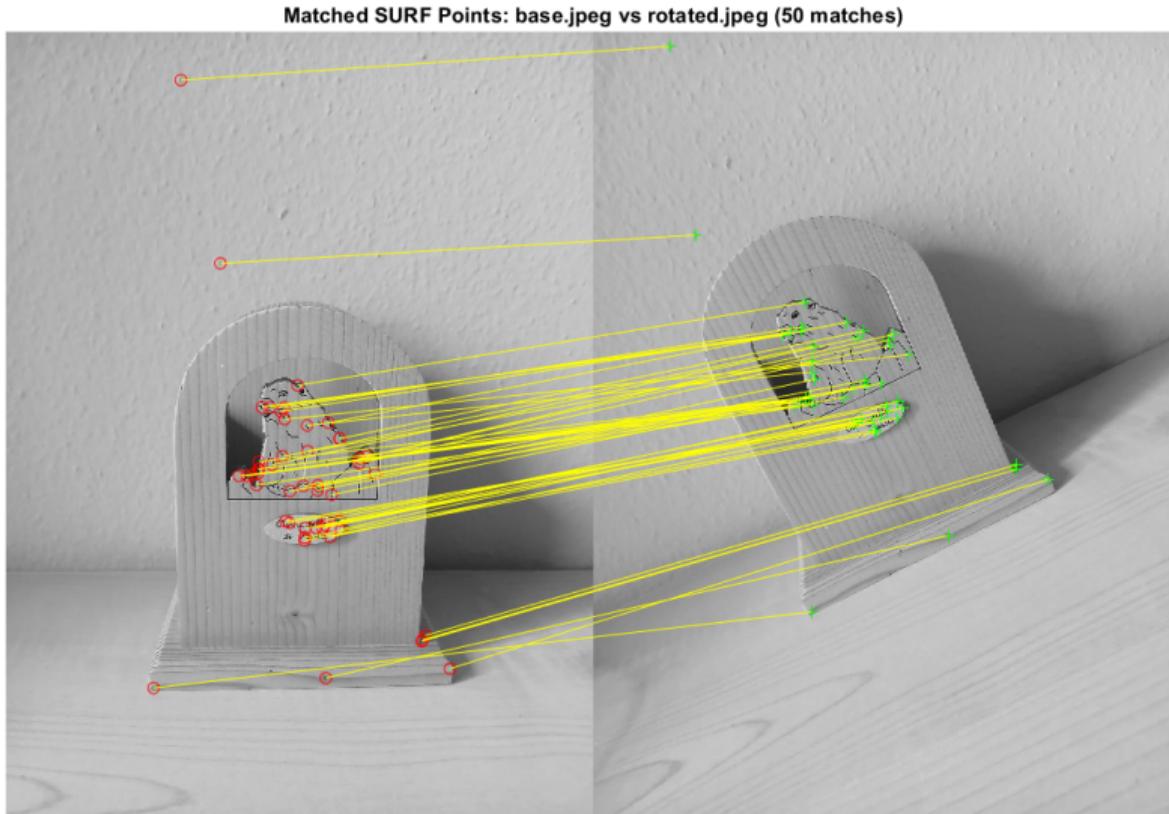
- Implementazione manuale: Immagine Integrale, Filtri Box, Soppressione Non-Massimale, formazione del descrittore (SURF).
- Pipeline SURF completa: Caricamento, Immagine Integrale, Rilevamento, Orientamento, Generazione descrittori, Corrispondenza.
- Il Feature Matching (corrispondenza) è stato realizzato con i software del pacchetto Vision Toolbox di Matlab.

Feature Matching

Matched SURF Points: base.jpeg vs close_up.jpeg (105 matches)



Feature Matching



Conclusione

- **Efficienza e Robustezza di SURF:** Comparabile a SIFT, ma più efficiente.
- **Innovazioni Chiave:** Immagini integrali, rilevamento Hessiano, descrittori Haar wavelet.
- **Feature Matching:** Fornisce un robusto vocabolario visivo per l'identificazione e l'abbinamento.
- **Validazione Empirica:** Negli esempi portati l'algoritmo è risultato robusto.
- **Utilità Pratica:** Cruciale per applicazioni di visione artificiale in tempo reale.

Lavori Futuri e Domande

- **Direzioni Potenziali:**

- Confronto con altri rilevatori (ORB, AKAZE).
- Ottimizzazione parametri e accelerazione GPU.
- Adattamento per domini specifici (imaging medico, 3D).
- Analisi più rigorosa degli errori di approssimazione.

Grazie! Domande? leonardo.fusarbassini@studenti.unimi.it