

Sistemi Intelligenti Algoritmi genetici

Alberto Borghese

Università degli Studi di Milano
Laboratory of Applied Intelligent Systems (AIS-Lab)
Dipartimento di Informatica
alberto.borghese@unimi.it



A.A. 2023-2024

1/76

<http://borghese.di.unimi.it/>



Sommario



- **Algoritmi genetici**
- Applicazione alla genetic art
- Ottimizzazione evolutivistica
- Applicazione delle strategie evolutive alla visione

A.A. 2023-2024

2/76

<http://borghese.di.unimi.it/>



Utilizzo degli algoritmi genetici



- Algoritmi di ottimizzazione (funzioni costo non esprimibili in modo analitico, valutazioni parziali...)
- Costruzione di dati a partire da un data-set di partenza, mescolando caratteristiche o sotto-insiemi di dati.



Evoluzione



- Processo discontinuo su un arco tempo reale ampio e su un numero di individui grande, che porta la specie al successo.
- **Algoritmi genetici.** Ottimizzazione evolutiva. Programmazione Genetica.
- **Computazione evolutivistica** è un termine generico che indica una gamma di sistemi di risoluzione di problemi di ottimizzazione che emulano l'evoluzione naturale nel ricercare la soluzione migliore.



Alcune pietre miliari



- 1960: Ingo Rechenberg introduce l'idea della computazione evolutivistica nel suo lavoro "Evolution strategies"
- L. Fogel 1962 (San Diego, CA): Programmazione evolutivistica
- I.Rechenberg, H.P.Schwefel, 1972, Ottimizzazione evolutiva.
- 1975: John Holland inventa gli algoritmi genetici e pubblica il libro "Adaption in Natural and Artificial Systems"
- 1992: John Koza propone gli algoritmi genetici per fare evolvere programmi software che effettuano specifici compiti. Tale metodo viene chiamato da lui Programmazione Genetica

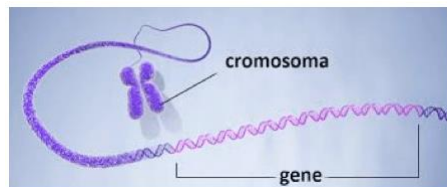


Struttura

Algoritmi Genetici



- **Cromosoma:** La codifica di una possibile **soluzione** per un dato problema di solito rappresentata tramite un array di bit o caratteri. Un cromosoma può essere associato ad una variabile, più variabili possono essere associate a diversi cromosomi.
- Corpuscolo all'interno della cellula contenente un tratto di DNA. Il DNA è una struttura allungata, a doppia elica, costituita da filamenti di basi affacciate e connesse a coppie: Adenina-Timina, Citosina-Guanina o Uracile.
- **Gene.** Un singolo bit o un insieme di bit, **parte della soluzione** (e.g. parte di una variabile).
- Sequenza di basi del DNA che codifica un'informazione (e.g. colore occhi), sottoparti di un cromosoma.



Funzionamento

- **Ricombinazione.** Vengono generate nuove soluzioni mescolando due soluzioni (cromosomi).
- Nel generare il cromosoma di un figlio, tratti omologhi del cromosoma dei genitori si scambiano materiale genetico.
- **Mutazione.** Cambiamenti casuali nella soluzione (nel cromosoma).
- "Errore" che avviene durante la replicazione del DNA generato dalla ricombinazione.



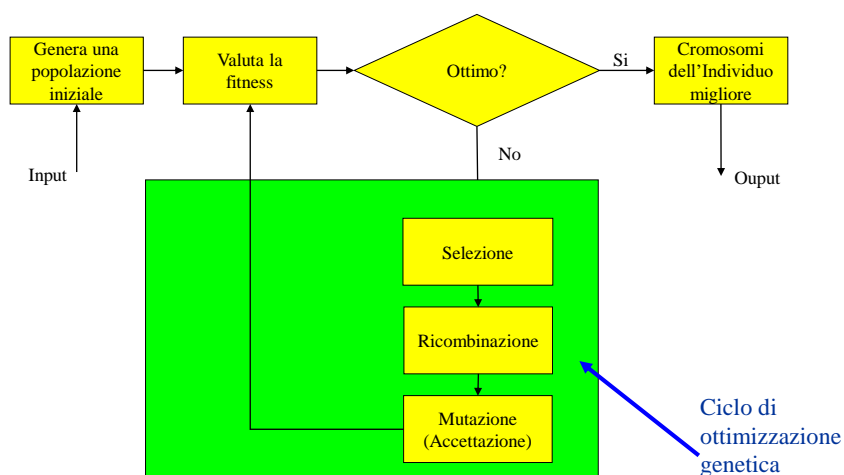
Ottimizzazione genetica



- **Codifica** della potenziale soluzione di uno specifico problema in una struttura dati associabile ad un cromosoma artificiale.
- **Evoluzione genetica.** Applicazione iterativa del processo di selezione, ricombinazione e mutazione a queste strutture.
- Valutazione del risultato mediante funzione di **fitness**
- **Fitness:** misura del successo di un individuo della specie in relazione alla sua **interazione con l'ambiente**.
- La fitness è misurata nell'interazione con l'ambiente (cf. reward).
- Non si riesce spesso a calcolare un gradiente della fitness, è spesso il risultato di un processo di valutazione.



Algoritmo genetico - flow chart



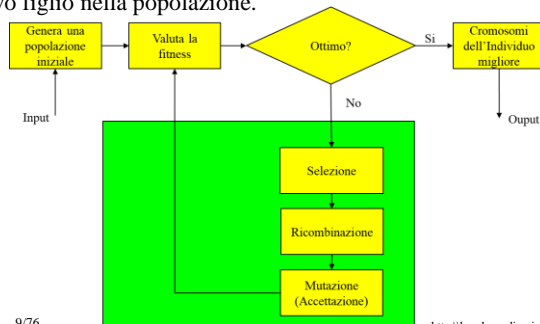


Algoritmo genetico - i passi



1. **[Inizializzazione]** Genera una popolazione random di n cromosomi (soluzioni plausibili - genitori)
2. **[Fitness]** Valuta la fitness $f(x)$ di ciascun cromosoma x nella popolazione corrente (bontà della soluzione associata al cromosoma).
3. **[Nuova popolazione]** Crea una nuova popolazione (prole) attraverso:
 - (a) **[Selezione]** Seleziona i cromosomi genitori in base alla loro fitness (criterio)
 - (b) **[Ricombinazione]** Con una certa probabilità di ricombinazione incrocia due genitori per generare un nuovo figlio.
 - (c) **[Mutazione]** Con una certa probabilità di mutazione cambia i geni che costituiscono il figlio
 - (d) **[Accettazione]** Introduce il nuovo figlio nella popolazione.

La soluzione è ricavata dall'elemento con la miglior fitness.



A.A. 2023-2024

9/76

<http://borghese.di.unimi.it/>



Gli elementi da definire



Codifica della soluzione

- Come creare i cromosomi e che tipo di codifica utilizzare

Ciclo di ottimizzazione genetica

- Come selezionare i genitori per la ricombinazione nella speranza che i genitori migliori produrranno migliore prole
- Come definire la ricombinazione.
- Come definire la mutazione.

A.A. 2023-2024

10/76

<http://borghese.di.unimi.it/>



Codifica della soluzione



La soluzione è codificata nel **cromosoma (insieme di geni)**.

- Un cromosoma deve contenere l'informazione sulla soluzione che rappresenta
- La codifica dipende principalmente dal problema da risolvere:
 - Codifica binaria implicita (numero binario -> soluzione)
 - Codifica binaria esplicita (numero binario = soluzione)
 - Codifica reale esplicita
 - Codifica mediante permutazione
 - Codifica mediante alberi di parsing....

Ne esaminiamo alcune.



Codifica binaria implicita



- La codifica binaria (logica) è la più comune negli algoritmi genetici.
- La **codifica implicita** (la prima ricerca di GA utilizzava tale codifica) prevede:
 - Codifica binaria: ogni cromosoma è una stringa di bit (1 o 0)
 - L'implementazione di operatori genetici risulta essere immediata
 - Non rappresenta la soluzione direttamente, ma come ricavarla

Esempio: problema dello zaino (cf. bioinformatica, logistica)

Date un certo numero di oggetti caratterizzati ciascuno da un valore e una dimensione, massimizzare il valore degli oggetti in uno zaino che ha una capacità fissata.

In questo caso, ogni bit può rappresentare il fatto che un oggetto sia inserito o meno nello zaino.

Cromosoma 1

1	1	0	1	1	0	0	1	0	0	1	1	0	1	1	0
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Cromosoma 2

1	1	1	0	1	1	1	0	0	0	0	1	1	1	1	0
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---



Codifica di permutazioni



La codifica di permutazione è una particolare codifica implicita. Essa viene utilizzata per problemi di ordinamento: il cromosoma è una stringa di numeri che rappresenta la posizione in una sequenza.



Esempio: problema del commesso viaggiatore (Traveling salesman problem)

È dato un insieme di città e le corrispondenti distanze a coppie. Il commesso viaggiatore deve visitarle tutte ma non vuole viaggiare più del necessario.

Scopo: Trovare una sequenza di città che minimizza la distanza percorsa.

Il cromosoma descrive l'ordine delle città visitate

Cromosoma 1

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

Cromosoma 2

2	9	1	4	1	1	1	5	8	1	5	1	3	6	1	2	1	6	7	3	1	0	4
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---



Codifica esplicita generalizzata



La codifica esplicita o diretta codifica direttamente la soluzione. Essa è indicata in problemi dove le entità coinvolte sono caratterizzate da una certa complessità

- Il cromosoma è una sequenza di geni, cioè di valori connessi al problema (numeri, caratteri, stringhe di caratteri, ...)
- Rappresenta la scelta migliore per ottimizzazione in spazi multidimensionali e/o multimodali. Gli operatori di ricombinazione e mutazione richiedono specifici controlli di consistenza

Esempio: addestramento di un agente (determinazione della policy)

I numeri reali nei geni del cromosoma rappresentano le probabilità di scegliere le azioni in un certo stato.

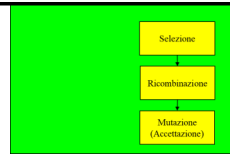
A	B	D	H	Y	V	S	V
---	---	---	---	---	---	---	---

2.5678	1.4361	3.3426	7.8761
--------	--------	--------	--------

open	walk	back	close
------	------	------	-------



Selezione



- In accordo alla teoria evuzionistica di Darwin il migliore cromosoma sopravvive per creare nuova prole. Per selezionare il migliore cromosoma esistono diversi criteri:
 - Ordinamento (priorità diretta ai cromosomi che si sono espressi meglio)
 - Casualità (eventualmente solo per i cromosomi migliori)
 - ...

In particolare è possibile dare una probabilità di selezione crescente proporzionale alla fitness (**Roulette**)



Selezione tramite roulette



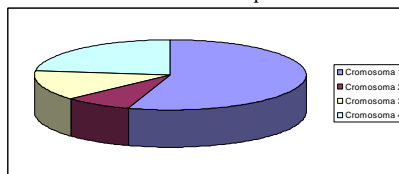
Gli individui sono selezionati proporzionalmente alla loro fitness. Migliore essa è e più alta è la probabilità di selezione

1. Si immagina una roulette dove sono sistemati tutti i cromosomi della popolazione
2. La dimensione della sezione nella roulette è proporzionale al valore di fitness di ciascun cromosoma
3. La pallina viene lanciata all'interno della roulette e il cromosoma in corrispondenza del quale si ferma è quello selezionato

Esempio:

Generazione casuale di 6 numeri (0 1)

0.81, 0.32, 0.96, 0.01, 0.65, 0.42

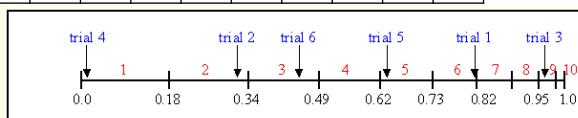


Numero di individui	1	2	3	4	5	6	7	8	9	10	11
Fitness	3.0	2.7	2.4	2.1	1.8	1.5	1.2	0.9	0.6	0.3	0.0
Probabilità di selezione	0.18	0.16	0.15	0.13	0.11	0.09	0.07	0.06	0.03	0.02	0.00

$$\sum_k f_k = 16.5$$

Cromosoma 1:

$$P(\text{figlio} = 1) = 3.0 / 16.5 = 0.1818$$



Seleziono gli elementi 1, 2, 3, 5, 6, 9 per la generazione successiva

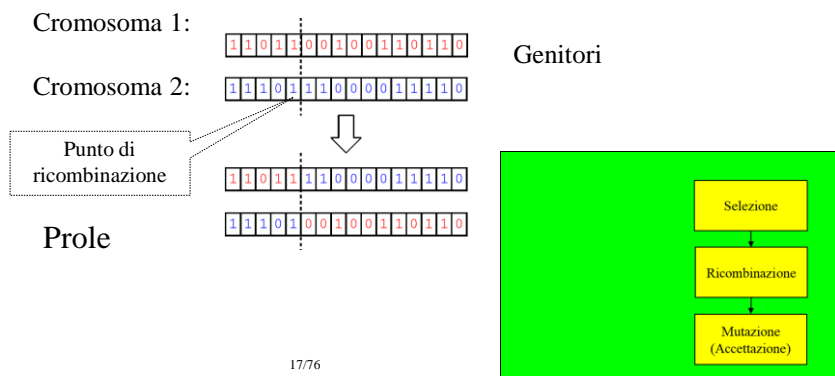


Ricombinazione (binaria)



Il processo di ricombinazione di base opera su geni selezionati dal cromosoma genitore e crea nuova prole:

1. Selezione random di un **punto di ricombinazione** all'interno del cromosoma
2. Copia tutti i geni precedenti questo punto dal primo genitore e poi copia tutti i geni successivi a questo punto dal cromosoma del secondo genitore



A.A. 2023-2024

17/76



Estensione della ricombinazione binaria



Per la codifica binaria esistono diverse modalità:

- **Ricombinazione da un singolo punto:** selezionato un punto all'interno del cromosoma, la stringa binaria compresa tra l'inizio e il punto di ricombinazione è copiata dal primo genitore, il resto è copiato dal secondo genitore
- **Ricombinazione da due punti:** selezionati i due punti all'interno del cromosoma, la stringa binaria compresa tra l'inizio e il primo punto di ricombinazione è copiata dal primo genitore, la parte compresa tra il primo e il secondo punto è copiata dal secondo genitore, la parte compresa tra il secondo punto e la fine del cromosoma è copiata ancora dal primo genitore
- **Ricombinazione uniforme:** i bit sono copiati casualmente dal primo o dal secondo genitore
- **Ricombinazione aritmetica:** operatori algebrici (e.g., logica AND, OR, ..). Altre funzioni booleane possono essere utilizzate.

Nella ricombinazione a più punti si possono considerare più genitori.

In alcuni problemi la ricombinazione può generare soluzioni non ammissibili. Queste soluzioni non evolveranno ulteriormente.

A.A. 2023-2024

18/76

<http://borgnese.di.unimi.it/>



Mutazione (binaria)

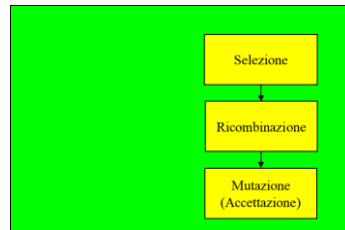
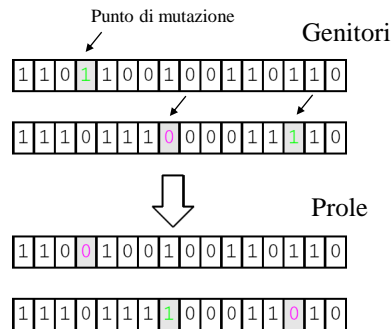


Dopo la ricombinazione interviene il processo di mutazione

Nella mutazione di base cambia in modo casuale un numero contenuto di bit da 0 a 1 o viceversa da 1 a 0

Problema:

Come conservare la consistenza della soluzione (e.g. problema del commesso viaggiatore)?



A.A. 2023-2024

19/76



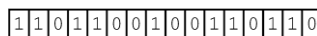
Elitismo



Come preservare i cromosomi che danno soluzioni già buone?

- Quando si selezionano i nuovi padri si tiene conto della loro fitness, implicitamente si salvano via via le soluzioni migliori.
- Nel caso della codifica binaria si può introdurre un “blocco” o una “resistenza” alla modifica dei geni migliori.

blocco →



A.A. 2023-2024

20/76

<http://borgese.di.unimi.it/>



Riflessioni



Che relazione c'è tra i metodi di selezione utilizzati negli algoritmi genetici e le politiche ϵ -greedy di scelta dell'azione?

Che differenze e similitudini ci sono con gli algoritmi di ottimizzazione classici (cf. algoritmo del gradiente)?

Come si può fare per estrarre più soluzioni lungo le direzioni dello spazio delle soluzioni lungo le quali la funzione obiettivo diminuisce più rapidamente?

=

Come si possono scegliere più individui lungo le direzioni in cui la fitness aumenta più rapidamente?

Introduzione di variabili *strategiche*.



Sommario



- Algoritmi genetici
- **Applicazione alla genetic art**
- Ottimizzazione evolutiva
- Applicazione delle strategie evolutive alla visione
- Aggiornamento dei parametri strategici



Emotional Intelligence



Cosa non può fare una macchina?

“Essere gentile, bellissima, amichevole, avere iniziativa, essere spiritosa, distinguere il bene dal male, innamorarsi, essere golosa di fragole alla panna, fare innamorare qualcuno, imparare dall’esperienza, avere proprietà di linguaggio, riflettere su sé stessa, avere una varietà di comportamenti, essere veramente creativa” (A. Turing, 1940).

A. Damasio, *Descartes’ error: Emotion, Reason, and the Human Brain*, Putnam, 1994.

Goal 1: computer systems capable of making normative judgements related to questions of beauty and taste in the arts.

Goal 2: meta-aesthetic exploration of emerging aesthetics in machines.



A.A. 2023-2024

23/76



Painting



How do we evaluate a painting? Which emotions arise?

How do painters come up with micro-decision during painting (change color, angle...)

How do they get new «knowledge» from the evaluation of terminated paintings?

How does the circular reaction (through mirror neurons) work for perception and execution of painting? Does it apply to painting?

Which is the neural substrate of aesthetic judgement (Neuroesthetics)?

Largely Unknown

A.A. 2023-2024

24/76

<http://borghese.di.unimi.it/>



Evaluation metric (fitness)

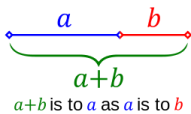


Birhoff's Aesthetic Measure (1933). $M = O / C$

O = degree of order
C = degree of complexity

Zipf's law (Manaris et al., 2005). $P_i = 1 / \text{rank}_i^a$ with a close to 1.

Golden ratio $\varphi = 1 + \frac{1}{\varphi}$ $\varphi = \frac{1 + \sqrt{5}}{2} = 1.6180339887\dots$

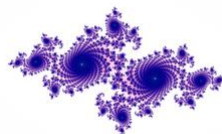


Euclid (300 BC)
Livio (2003)



Fractals (Paul Jackson Pollock painting - Taylor 2006)

Forme che si ripetono su scale diverse.



A.A. 2023-2024



Computational methods have been tried (Datta et al., 2006)



3851 Pictures from a Photo.net scored, peer-rated [1-7], in terms of Aesthetics and originality.



Binary classifier (SVM with 5-fold cross-validation) on 56 features.

Features extracted from HSV images according to experience of photographers.

- 1/3 rule (approximation of the golden ratio)
- Colorfulness, texture richness, Depth Of Field...
- Familiarity (with respect to a data base defined for image retrieval....)

Tree-based (hierarchical) clustering to understand the importance of the different features.

A.A. 2023-2024

26/76

<http://borghese.di.unimi.it/>

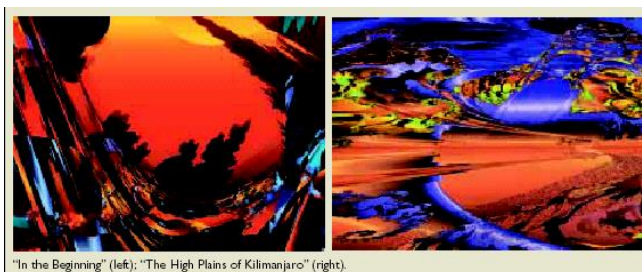


Karl Sims's pioneer work in genetic art



Siggraph 2001 and afterwards

Geometric patterns



"In the Beginning" (left); "The High Plains of Kilimanjaro" (right).



The genotype characteristics: frequency



Genes are defined here as basic functions. Which is their proper scale?

Trigonometric functions with $\omega = 1$

```

sin(x*x + y*y)
sin(x*x)*cos(y*y)
sin(x/y)*cos(z/y)
cos(x/y)
sin(y/x)
abs(y)-x
x-abs(y)
abs(x)
abs(y)
abs(x)*abs(y)
sin(x)*cos(y)
sin(x*y)*cos(x*y)
sin(x*x-y*y)
sin(x*x)
y-abs(x)
y-sin(x)
x-cos(y)
abs(x)+y
sin(x*x-x*y*y*y)
sin(y*y*y)+sin(x*x*x)
cos(y*y*y)+x*x*x
cos(y*y*y)+cos(x*x*x)
abs(y*3)
abs(x*3)
sin(x*x/y-y*y/x)

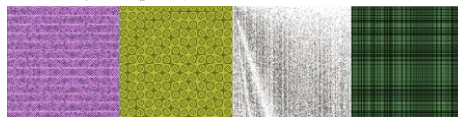
```

```

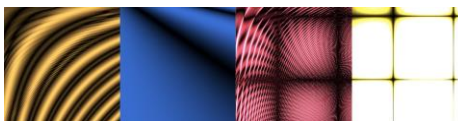
cos(x*x/y)+sin(y*y/x)
sin(x)+sin(x)+cos(y)+cos(y)
cos(x)+cos(x)+sin(y)+sin(y)
sin(x)+cos(x)+sin(y)+cos(y)
cos(y)+sin(y)+cos(x)+sin(x)
tan(cos(sqrt(x*y*x*y)))
sqrt(abs(x)+abs(y))
sin(x*y*x)+cos(y*x*x*y)
sin(sqrt(abs(x)))-cos(sqrt(abs(y)))
sqrt(cos(x)+sqrt(x)*sin(y)+sqrt(y))
cos(x)*sin(x*y)
cos(y)*sin(x*y)
sin(x+y*x*y+x*x)
sin(y+x*y*x+y*y)
abs(x*y+x*x*y*y)
((x+y)*y*x*sin(x)*cos(y))
((x-y)*x)+sin(x*y)+cos(y/x)
sin(x*y)+cos(y*x+y)
cos(x+y)*sin(x+y)/2
cos(sqrt(x+y))*y+sqrt(cos(y)*sin(x))
sin(sqrt(y-x))*x+sqrt(sin(x)*cos(y))
cos(x)*sin(x)-cos(y)*sin(y)
sin(abs(cos(x+y))+abs(cos(y*x*y)))
sin(cos(x)*abs(y)*abs(y))
cos(x)*sin(y)*cos(x*y)

```

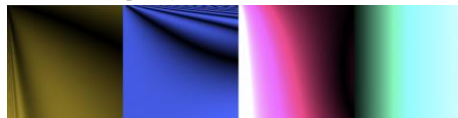
High Frequency (hx with $h = 1, T = 2\pi$)



Medium Frequency (hx with $h = N/6, T = 2\pi(N/6)$)



Low Frequency (hx with $h = N, T = 2\pi N$)



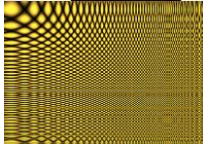


The phenotype characteristics: intensity



[R, G, B] images. Which palette for each channel? At each generation, random level for each color.

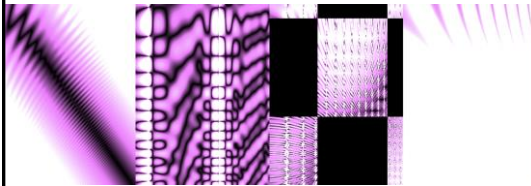
R = 200, B = 151, G = 21



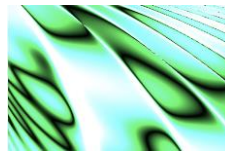
R = G = B = 128



R = B = 255; G = 100



R = 47, G = 252, B = 10



A.A. 2023-2024

29/76

<http://borghese.di.unimi.it/>



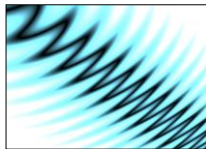
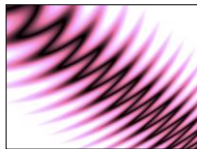
Mutation and Recombination



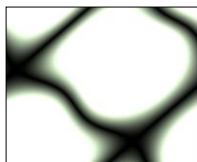
No recombination of several images has been considered.

The different subfunctions are combined through arithmetic operators: {+, -, *, /}

Each image is mutated by changing 1 function and 1 operator, randomly chosen.



2 successive generations
low difference



2 successive generations
large difference

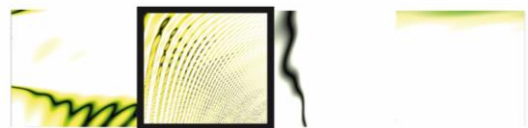
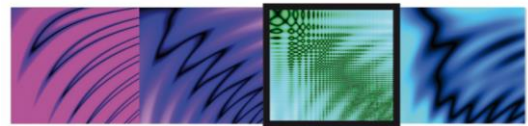
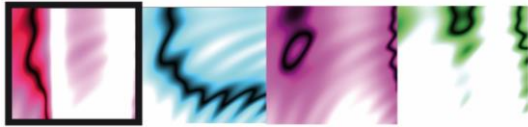
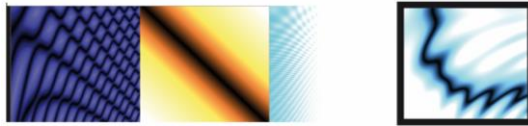
A.A. 2023-2024

30/76

<http://borghese.di.unimi.it/>



Sequence of generations



More info in:
R. Bellini, N.A. Borghese
(2013) Genetic arts in
perspective. *Proc. XXI
WIRN conference.*
2013.

Collective
Evaluation

WEB polls

A.A. 2023-2024

31/76

<http://borghese.di.unimi.it/>



Evolution



$\{\lambda, \mu\}$ evolutionary optimization. λ parents, μ sons for each parent.

The best son can be chosen or the best between sons and parents. How?

Fractal dimension? Golden Section? Zipf's law?.... Evaluation by an expert!



©OriginalArtist
Reproduction rights obtainable from
www.CartoonStock.com

"Too bright?"

A.A. 2023-2024

nimi.it/



The next Rembrandt project



- ING, Microsoft, Delft University of Technology, The Mauritshuis and Museum Het Rembrandthuis (<https://www.nextrembrandt.com/>).
- The AI, capable of deep learning, was imprinted with 346 of Rembrandt's known works in the hopes that it can create a unique 3D printed image in his style.

- Identification of features of faces (e.g. nose, eyes, chin...) and probability.
- Geometrical transformation of features (rotation / scaling / translation)
- Assembling features and validating them according to proportions.
- Lighting



Can we build artificial painters?

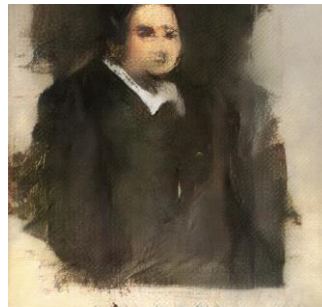
A.A. 2023-2024

33/76

<http://borghese.di.unimi.it/>



GAN networks e pittura (I. Goodfellows)



Ritratto di Edmond Belamy Creato dal collettivo francese Obvious e Battuta da Christie nel 2018

A.A. 2023-2024

34/76

<http://borghese.di.unimi.it/>



Some open questions



Is there any evolution path? Is there any «optimum»?

On which features or parameters is evolution based?

Which are the determinant of aesthetic evaluation?

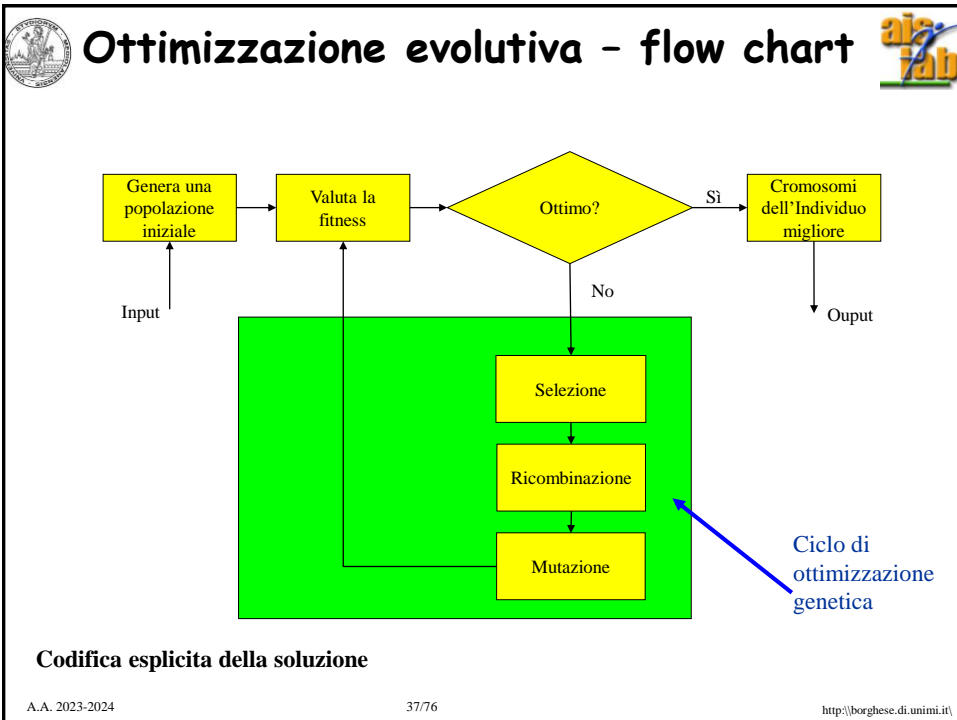
Is there any particular component (or gene) that is more important than the others? Is there any form of elitism?



Sommario



- Algoritmi genetici
- Applicazione alla genetic art
- **Ottimizzazione evolutiva**
- Applicazione delle strategie evolutive alla visione



I genitori della prima generazione

- I genitori iniziali sono scelti da una distribuzione statistica, solitamente uniforme
- Obiettivo è utilizzare la storia della popolazione per indirizzarne l'evoluzione.
- La selezione dei genitori nei modelli più semplici di ES non è polarizzata: ciascun individuo ha la stessa probabilità di essere scelto
- Tra una generazione e la successiva avvengono gli eventi evolutivi:
 - Selezione
 - Ricombinazione
 - Mutazione

A small version of the flowchart from the previous slide is shown at the bottom right of the slide.

A.A. 2023-2024 38/76 <http://borghese.di.unimi.it/>



Selezione in ES



μ - # genitori λ - # figli.

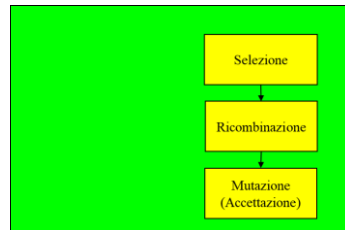
Metto in competizione padri e figli nel processo di selezione (individui della generazione precedente e individui della generazione attuale)

Strategia $(\mu + \lambda)$ – I genitori della generazione successiva vengono scelti tra i $\mu + \lambda$ elementi della popolazione attuale. E' una strategia elitista.

Strategia (μ, λ) – I genitori della generazione successiva vengono scelti solamente tra i λ figli. E' la strategia preferita, soprattutto con adattamento della matrice di covarianza. La strategia $(\mu + \lambda)$ tende a concentrare le soluzioni in luoghi che possono dare soluzioni sub-ottimali.

$\lambda \gg \mu$ e.g. $\lambda = 7 \mu$

Gli elementi selezionati vengono poi ricombinati per diventare i genitori della generazione successiva.



A.A. 2023-2024

39/76



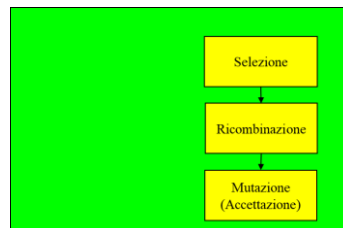
Ricombinazione in ES



Dato un certo numero di figli, come saranno fatti i genitori della generazione successiva?

- Potrei definire come genitori, gli elementi migliori (massima fitness) della generazione precedente.
- Potrei scegliere come genitori (con probabilità più bassa) elementi meno buoni che possono portare ad esplorare nuovi spazi di soluzione (cf. metodo della roulette). Anche gli individui valutati meno buoni possono portare ad una stirpe migliore.....
- **Potrei scegliere come genitori un mix di diversi individui**

□ **Elitismo.** Quando si selezionano i nuovi padri si tiene conto della loro fitness, implicitamente si salvano via via le soluzioni migliori.



A.A. 2023-2024

40/76

<http://borgese.di.unimi.it/>



Ricombinazione in ES



Agisce sui figli correnti per determinare i nuovi genitori:

- Combinando cromosomi corrispondenti agli elementi (nel caso comune di più elementi)

- Ricombinazione discreta (e.g. colore RGB di un pixel)**

Individuo 1: [12 25 5]	estrazione 1: 2 2 1	Figlio 1: [123 4 5]
Individuo 2: [123 4 34]	estrazione 2: 1 2 3	Figlio 2: [12 4 45]
Individuo 3: [7 27 45]	estrazione 3: 1 2 2	Figlio 3: [12 4 34]

- Ricombinazione intermedia**

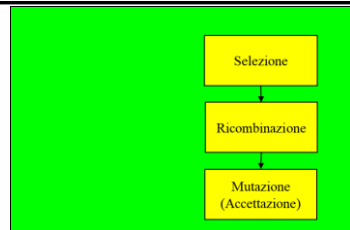
Figlio = individuo1 + α (individuo2-individuo1) $0 \leq \alpha \leq 1$
 (estendibile alla combinazione lineare di n individui)

Il figlio generato si trova all'interno del convex hull dei padri (per uscire occorrerebbe $\alpha > 1$, ma la generazione di figli esterni al convex hull e' demandata alla mutazione).

- Selezionando uno tra gli elementi disponibili.
 - Esempio di soluzione a valori discreti → elemento intermedio
 - Esempio di soluzione a valori reali: → posizione media.



Mutazione in ES



Elemento^{t+1} = Elemento^t + z

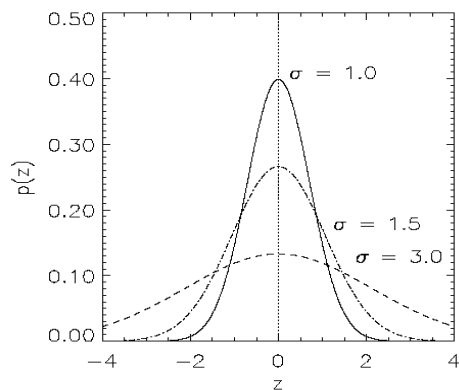
Mutazione: perturbare i valori aggiungendo un rumore casuale Spesso è utilizzata una distribuzione Gaussiana a media nulla e deviazione standard σ ,

$N(0, \sigma) \rightarrow x'_i = x_i + N(\sigma, 0)$

- Media nulla
- Deviazione standard σ

La deviazione standard σ è chiamata **passo di mutazione** ed è una proprietà comune a **tutti i geni**.

σ definisce l'ampiezza dello spazio locale di ricerca (isotropo lungo tutte le variabili - geni).



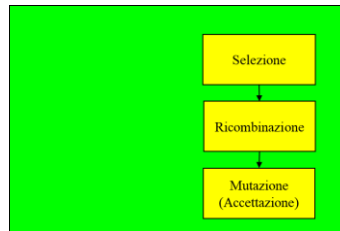


Adattamento della mutazione



L'ampiezza della mutazione consente di definire attorno alla soluzione associata al figlio una regione all'interno della quale la nuova soluzione associata al figlio mutato può cadere.

- 1/5 strategy
- Matrice di covarianza



Adattamento dell'ampiezza della regione di mutazione - I



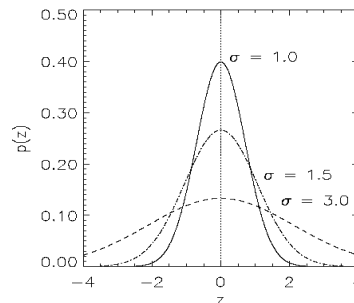
$$\mathbf{x}'(n) = \mathbf{x}(n) + \sigma'(n) \cdot \mathbf{N}(0,1)$$

- σ viene aggiornata run-time con la regola euristica "1/5 success rule" (I. Rechenberg. *Evolution strategies. 1973* → Doerr & Doerr, *GECCO 2015*).

Questa regola ricalcola σ ogni k epoche tramite $r(n)$:

$$\begin{aligned} \rightarrow r(n) &= 1/c > 1 && \text{se } p_s > 1/5 \\ \rightarrow r(n) &= c < 1 && \text{se } p_s < 1/5 \\ \rightarrow r(n) &= 1 && \text{se } p_s = 1/5 \end{aligned}$$

dove p_s è la percentuale di mutazioni con successo, $0.8 \leq c < 1$.



Smaller steps occur more likely than big ones, until a child performs better than its ancestor and takes its place.

$$\sigma'(n) = \sigma(n) \cdot r(n) = \mathbf{RF}(\sigma)$$

$$\mathbf{x}'(n) = \mathbf{x}(n) + \mathbf{RF}(\sigma) \mathbf{N}(0,1) = \mathbf{x}(n) + \mathbf{N}(0,\sigma)$$



Esempio di mutazione per un singolo gene



Valuto se il figlio è migliore del padre. Supponiamo ES (1+10)

- Es: $k=10$
 - $k=1$ -- fitness figlio > fitness padre
 - $k=2$ -- fitness figlio > fitness padre
 - $k=3$ -- fitness figlio < fitness padre
 - $k=4$ -- fitness figlio > fitness padre
 - $k=5$ -- fitness figlio < fitness padre
 - $k=6$ -- fitness figlio > fitness padre
 - $k=7$ -- fitness figlio < fitness padre
 - $k=8$ -- fitness figlio < fitness padre
 - $k=9$ -- fitness figlio > fitness padre
 - $k=10$ -- fitness figlio > fitness padre

- $6/10 > 1/5$ quindi σ cresce $\sigma = \sigma / c$.

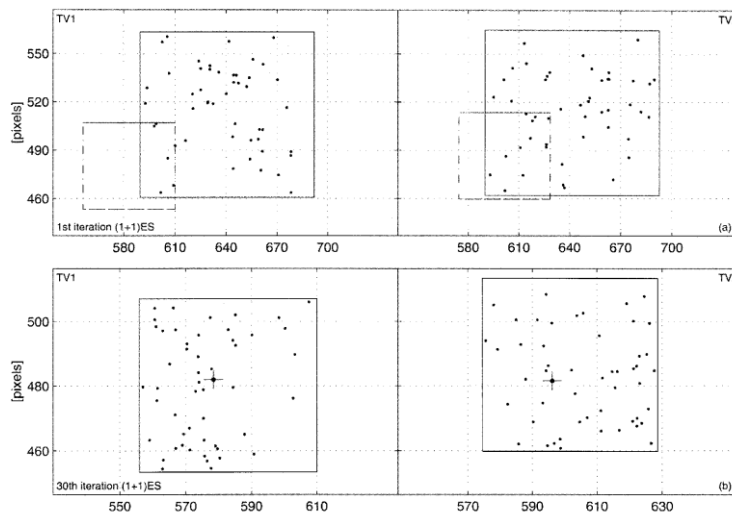
σ decresce esponenzialmente ad ogni generazione. In aggiunta, se $p_s \ll 1/5$, si somma uno scalino crescente o decrescente.



Esempio



Si parte da una regione ampia centrata in $\{640, 640\}$. Dopo un certo numero di iterazioni si arriva ad una regione di ricerca più piccola attorno a $\{581, 482\}$ per la prima camera e $\{600, 490\}$ per la seconda camera.





Adattamento della mutazione avanzato



Vorrei trovare un modo di adattare l'ampiezza della regione di ricerca nelle diverse direzioni (variabili) in funzione della fitness (bontà) della soluzione.

*Correlazione tra fitness e variabili
(figli nelle varie generazioni)*

Introduzione di variabili di stato



Variabili strategiche



Ogni gene di un cromosoma (variabile) ha associato **un'ampiezza di mutazione personalizzata**, mutazione scorrelata.

La **mutazione** dipende da σ :

$$\text{Figlio}^t = \text{Elemento}^t + z, \text{ con } z = N(0, \sigma) \rightarrow \mathbf{x}^{(g+1)}(n) = \mathbf{x}^{(g)}(n) + \mathbf{z}$$

Quanto vale σ ?

Cromosomi: $[x_1, \dots, x_n; \mathbf{w}]$

$\{x_1, x_n\}$ rappresentano gli elementi della soluzione (ad esempio gli elementi di un vettore), $\{\mathbf{w}\}$ rappresentano le **variabili strategiche**, ad esempio σ , che determinano l'ampiezza della mutazione.

$$\mathbf{x}'(n) = \mathbf{x}(n) + \mathbf{N}(0, \sigma) = \mathbf{x}(n) + \begin{bmatrix} \sigma & 0 \\ 0 & \sigma \end{bmatrix} \begin{bmatrix} N(0,1) \\ N(0,1) \end{bmatrix}$$



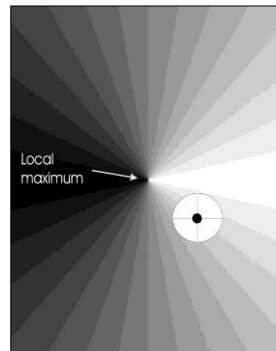
Spazio isotropo di mutazione



La mutazione **isotropa** genera figli distribuiti uniformemente all'interno di un'ipersfera nello spazio delle soluzioni

$$\mathbf{x}'(n) = \mathbf{x}(n) + \sigma \mathbf{N}(0,1)$$

Ampiezza della regione: σ



$$\mathbf{x}'(n) = \mathbf{x}(n) + \mathbf{N}(0,\sigma) = \mathbf{x}(n) + \begin{bmatrix} \sigma & 0 \\ 0 & \sigma \end{bmatrix} \begin{bmatrix} N(0,1) \\ N(0,1) \end{bmatrix}$$



Mutazione - σ personalizzato



- Inserisco valori diversi di σ per diverse variabili, cioè per i diversi cromosomi. Concedo che lo spazio delle soluzioni sia esplorato maggiormente lungo alcune direzioni rispetto altre.
- Mutazione scorrelata con un valore di σ_i associato ad ogni cromosoma.
 - Cromosomi: $[x_1, \dots, x_n, \sigma_1, \dots, \sigma_n] = [x_1, \dots, x_n; \mathbf{w}]$
 - $x'_i(n) = x_i(n) + \sigma'_i(n) \cdot N_i(0,1)$ alla generazione n

In forma vettoriale / matriciale:

$$\square \mathbf{x}'(n) = \mathbf{x}(n) + \mathbf{D}(n) \cdot \mathbf{N}_i(0,1)$$

Come aggiornare in modo diverso l'ampiezza della mutazione associata ai diversi cromosomi?

$$\mathbf{x}'(n) = \mathbf{x}(n) + \mathbf{N}(0,\mathbf{D}) = \mathbf{x}(n) + \begin{bmatrix} \sigma_{11} & 0 \\ 0 & \sigma_{22} \end{bmatrix} \begin{bmatrix} N(0,1) \\ N(0,1) \end{bmatrix} = \begin{bmatrix} D_{11} & 0 \\ 0 & D_{22} \end{bmatrix} \begin{bmatrix} N(0,1) \\ N(0,1) \end{bmatrix}$$

$$\mathbf{w} = \{D_{11}, D_{22}\}$$



Spazio anisotropo di mutazione



La mutazione isotropa genera figli distribuiti uniformemente all'interno di un'ipersfera nello spazio delle soluzioni

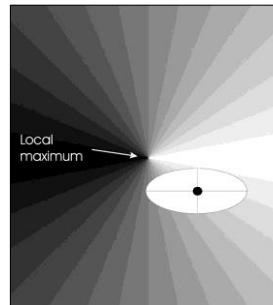
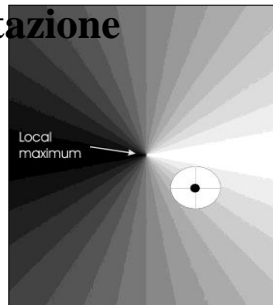
$$\mathbf{x}'(n) = \mathbf{x}(n) + \sigma \mathbf{N}(0,1)$$

$$(\mathbf{x} = [x_1 \ x_2] \in \mathbb{R}^2)$$

La mutazione anisotropa genera figli distribuiti uniformemente all'interno di un'iperellissoide nello spazio delle soluzioni

$$\mathbf{x}'(n) = \mathbf{x}(n) + \mathbf{N}(0, \mathbf{D}) = \mathbf{x}(n) + \begin{bmatrix} D_{11} & 0 \\ 0 & D_{22} \end{bmatrix} \begin{bmatrix} N(0,1) \\ N(0,1) \end{bmatrix}$$

D^2 rappresenta la matrice di covarianza della generazione della popolazione.



Deformazione della regione di mutazione



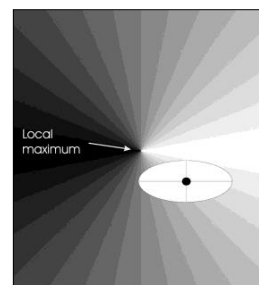
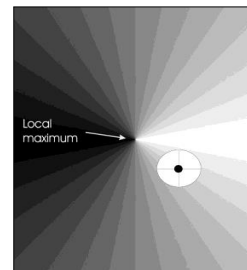
Supponiamo la superficie di equiprobabilità di mutare la soluzione si una circonferenza centrata nella soluzione attuale. La circonferenza potrà essere espressa come: σ^2

$$\mathbf{N}(0,1) = \mathbf{N}(0, \sigma^2) = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix} \begin{bmatrix} N(0,1) \\ N(0,1) \end{bmatrix}$$

Supponiamo la superficie di equiprobabilità di mutare la soluzione sia un'ellisse elongata come uno degli assi coordinati. L'ellisse potrà essere espressa come una Gaussiana isotropa per una matrice di deformazione: D

$$\mathbf{N}(0,1) = \mathbf{N}(0, D) = \begin{bmatrix} D_{11} & 0 \\ 0 & D_{22} \end{bmatrix} \begin{bmatrix} N(0,1) \\ N(0,1) \end{bmatrix}$$

$$D_{11} \neq D_{22}$$





Deformazione della regione di mutazione



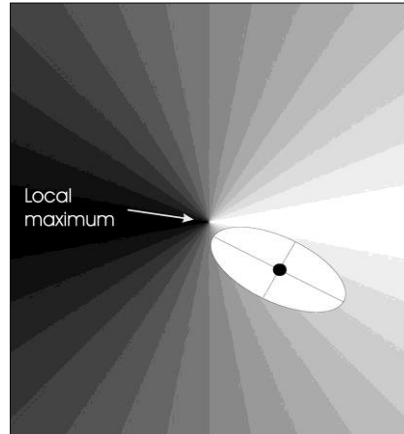
Supponiamo la superficie di equiprobabilità di mutare la soluzione sia un'ellisse elongata in direzione arbitraria nello spazio. L'ellisse potrà essere espressa come una Guassiana isotropa per una matrice di deformazione, \mathbf{D} , per una matrice di rotazione, \mathbf{B} :

$$\mathbf{B} \mathbf{D} \mathbf{N}(0, 1) = \mathbf{N}(0, \mathbf{B}\mathbf{D}) = \mathbf{N}(0, \mathbf{H})$$

dove $\mathbf{H} = \mathbf{B}\mathbf{D}$

La matrice $\mathbf{B}\mathbf{D}$ è legata alla matrice di covarianza, \mathbf{C} : $\mathbf{C}\mathbf{C}^T = (\mathbf{B}\mathbf{D})(\mathbf{B}\mathbf{D})^T = \mathbf{B} \mathbf{D}^2 \mathbf{B}^T$.

□ Cromosomi: $[x_1, \dots, x_n; \mathbf{B}, \mathbf{D}] = [x_1, \dots, x_n; \mathbf{w}]$



Rotazione della regione di mutazione



Come adattare la regione di ricerca al meglio?

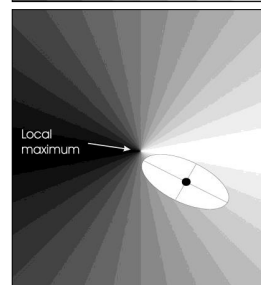
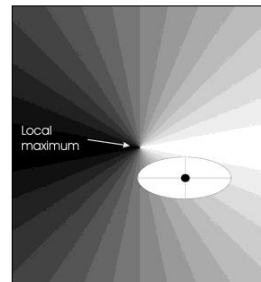
Ellissoide orientato nello spazio.

$$\Delta \mathbf{x}(n) = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} N(0,1) \\ N(0,1) \end{bmatrix} =$$

$$H_{12} = H_{21}$$

$$= \mathbf{B} \mathbf{D} \begin{bmatrix} N(0,1) \\ N(0,1) \end{bmatrix} \quad \mathbf{H} = \mathbf{B} \mathbf{D}$$

\mathbf{B} – matrice ortonormale (Matrice di rotazione)
 \mathbf{D} – matrice diagonale (asimmetria della lunghezza degli assi nelle due direzioni ortogonali orientate arbitrariamente)





Mutazione anisotropa



$$\mathbf{x}_k^{(g+1)} = \mathbf{x}_k^{(g)} + \rho^{(g)} \mathbf{H}^{(g)} \mathbf{z}_k^{(g)}$$

$$\mathbf{z}_k = [z_1 = N(0,1), z_2 = N(0,1), \dots, z_n = N(0,1)]^T \quad \text{Gaussiana}$$

- Vengono generati λ figli, \mathbf{x}_k , a partire dalla mutazione di un genitore, \mathbf{x}_j .
- $\mathbf{x}^{(g+1)}$ sarà funzione dei padri selezionati
- \mathbf{H} definisce l'iper-ellissoide, orientato arbitrariamente nello spazio, all'interno del quale effettuare le mutazioni a partire dal figlio generato.
- $\rho^{(g)}$ è un parametro di ampiezza globale che per ora consideriamo = 1.

Devo determinare H e per questo utilizzo l' "evolution path"

H ed evolution path rappresentano lo stato dell'evoluzione.



Evolution Path



La matrice \mathbf{D} viene aggiornata indirettamente tramite il cosiddetto **evolution path** \mathbf{s} .

$$\mathbf{s}^{(g+1)} = (1 - \alpha_s) \mathbf{s}^{(g)} + \alpha_u \mathbf{H}^{(g)} \mathbf{z}_{i_{sel}}^{(g)} \quad \mathbf{z}_{i_{sel}} \text{ rappresenta il genitore migliore}$$

$$\text{Nuovo_valore} = a * \text{Vecchio_valore} + b * \text{Innovazione}$$

Cf. Value estimate in RL

\mathbf{s} , evolution path, unisce tutti i genitori migliori. Dà un'idea dell'evoluzione. Path evolutivo all'interno dello spazio delle soluzioni. Connette i genitori migliori. \mathbf{s} viene rinforzato ogni volta che il genitore migliore si trova lungo l'evolution path.



Aggiornamento della matrice di Covarianza



$$\text{Nuovo_valore} = a * \text{Vecchio_valore} + b * \text{Innovazione}$$

$$\mathbf{H}^{(g+1)} = (1 - \alpha_C) \mathbf{H}^{(g)} + \alpha_C \mathbf{S}^{(g+1)}$$

$$\mathbf{S}^{(g+1)} = \mathbf{s} \mathbf{s}^T$$

Matrice di covarianza calcolata sul path evolutivo (s, mutazione dell'individuo migliore ad esempio)

Cosa vuol rappresentare questa equazione di aggiornamento?

Tanto più la covarianza rimane stabile, cioè la direzione e ampiezza di spostamento dell'individuo migliore rimane la stessa, tanto più \mathbf{H} cresce e si schiaccia in quella direzione → gli individui mutano (si spostano) più velocemente in quella direzione:

$$\mathbf{x}_k^{(g+1)} = \mathbf{x}_k^{(g)} + \mathbf{H}^{(g)} \mathbf{z}_k^{(g)}$$

A.A. 2023-2024

57/76

<http://borghese.di.unimi.it/>



Esempio -I



$$\mathbf{x}_k^{(g+1)} = \mathbf{x}_k^{(g)} + \rho^{(g)} \mathbf{H}^{(g)} \mathbf{z}_k^{(g)} \quad \rho^{(g)} = 1 \quad \mathbf{H}^{(g)} = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \quad \mathbf{s}^{(g)} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

$$\text{Si consideri il caso } \mathbf{z}^* = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \alpha_s = \alpha_c = 0.1 \text{ (poco peso all'aggiornamento)}$$

$$\mathbf{s}^{(g+1)} = (1 - \alpha_s) \begin{bmatrix} 2 \\ 0 \end{bmatrix} + \alpha_s \mathbf{H}^{(g)} \mathbf{z}^{*(g)} = \begin{bmatrix} 2.1 \\ 0 \end{bmatrix} \quad \mathbf{s} \text{ si allunga leggermente lungo } x \text{ (da 2 a 2.1)}$$

$$\mathbf{H}^{(g+1)} = (1 - \alpha_c) \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} + \alpha_c \begin{bmatrix} 4.41 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 3.141 & 0 \\ 0 & 0.9 \end{bmatrix} \quad \mathbf{H} \text{ si schiaccia e si allunga leggermente lungo } x$$

$$\mathbf{H}^{(g)} = \mathbf{B}^{(g)} \mathbf{D}^{(g)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\mathbf{H}^{(g+1)} = \mathbf{B}^{(g+1)} \mathbf{D}^{(g+1)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3.141 & 0 \\ 0 & 0.9 \end{bmatrix}$$

Modifica dell'ampiezza, \mathbf{D} , ma non della direzione \mathbf{B} .

A.A. 2023-2024

58/76

<http://borghese.di.unimi.it/>



Esempio - II



$$\mathbf{x}_k^{(g+1)} = \mathbf{x}_k^{(g)} + \rho^{(g)} \mathbf{H}^{(g)} \mathbf{z}_k^{(g)} \quad \rho^{(g)} = 1 \quad \mathbf{H}^{(g)} = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \quad \mathbf{s}^{(g)} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

Si consideri il caso $\mathbf{z}^* = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$ $\alpha_s = \alpha_c = 0.1$

$$\mathbf{s}^{(g+1)} = (1 - \alpha_s) \begin{bmatrix} 2 \\ 0 \end{bmatrix} + \alpha_s \mathbf{H}^{(g)} \mathbf{z}^{*(g)} = \begin{bmatrix} +1.5 \\ 0 \end{bmatrix} \quad \text{s si accorcia lungo x}$$

$$\mathbf{H}^{(g+1)} = (1 - \alpha_c) \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} + \alpha_c \begin{bmatrix} 2.25 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2.925 & 0 \\ 0 & 0.9 \end{bmatrix} \quad \mathbf{H} \text{ si contrae leggermente}$$



Esempio - III



$$\mathbf{x}_k^{(g+1)} = \mathbf{x}_k^{(g)} + \rho^{(g)} \mathbf{H}^{(g)} \mathbf{z}_k^{(g)} \quad \rho^{(g)} = 1 \quad \mathbf{H}^{(g)} = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \quad \mathbf{s}^{(g)} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

Si consideri il caso $\mathbf{z}^* = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ $\alpha_s = 0.1 \quad \alpha_c = 0.1$

$$\mathbf{s}^{(g+1)} = (1 - \alpha_s) \begin{bmatrix} 2 \\ 0 \end{bmatrix} + \alpha_s \mathbf{H}^{(g)} \mathbf{z}^{*(g)} = \begin{bmatrix} 1.8 \\ 0.1 \end{bmatrix} \quad \text{s ruota e si accorcia leggermente}$$

$$\mathbf{H}^{(g+1)} = (1 - \alpha_c) \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} + \alpha_c \begin{bmatrix} 3.24 & 0.18 \\ 0.18 & 0.01 \end{bmatrix} = \begin{bmatrix} 3.024 & 0.018 \\ 0.018 & 0.901 \end{bmatrix}$$

La matrice di covarianza non è più allineata con gli assi, vedremo più avanti come viene gestita. Si ha una crescita più contenuta dello schiacciamento.



Esempio - IV



$$\mathbf{x}_k^{(g+1)} = \mathbf{x}_k^{(g)} + \rho^{(g)} \mathbf{H}^{(g)} \mathbf{z}_k^{(g)} \quad \rho^{(g)} = 1 \quad D^{(g)} = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \quad \mathbf{s}^{(g)} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

Si consideri il caso $\mathbf{z}^* = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $\alpha_s = 0.1 \quad \alpha_c = 0.1$

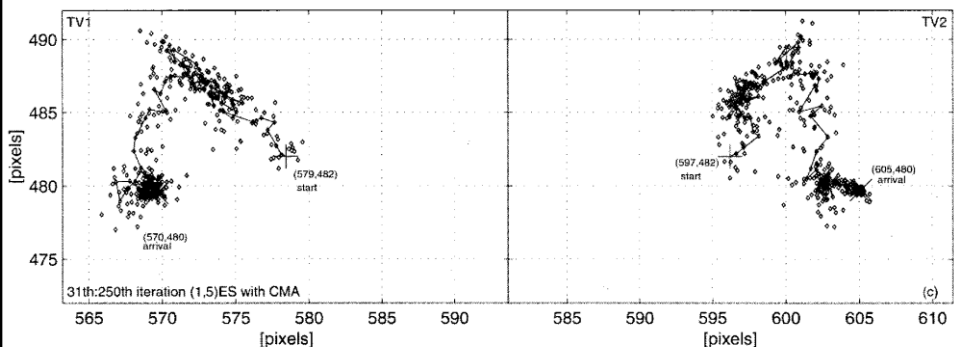
$$\mathbf{s}^{(g+1)} = (1 - \alpha_s) \begin{bmatrix} 2 \\ 0 \end{bmatrix} + \alpha_s \mathbf{H}^{(g)} \mathbf{z}^{*(g)} = \begin{bmatrix} 2.1 \\ 0.1 \end{bmatrix} \quad \text{s ruota e si allunga leggermente}$$

$$\mathbf{H}^{(g+1)} = (1 - \alpha_c) \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} + \alpha_c \begin{bmatrix} 4.41 & 0.21 \\ 0.21 & 0.01 \end{bmatrix} = \begin{bmatrix} 3.141 & 0.021 \\ 0.021 & 0.901 \end{bmatrix}$$

La matrice di covarianza non è più allineata con gli assi, vedremo più avanti come viene gestita. Si ha una crescita più contenuta dello schiacciamento, simile a quella ottenuta in precedenza. Si noti che in questo caso l'elemento migliore si sposta in diagonale ma con modulo $\sqrt{2}$ mentre nel caso precedente si spostava lungo y con modulo 1.



Evolution path





I parametri secondari



$$\mathbf{x}_k^{(g+1)} = \mathbf{x}_k^{(g)} + \rho^{(g)} \mathbf{B} \mathbf{D} \mathbf{z}_k$$

$$\mathbf{s}^{(g+1)} = (1 - \alpha_s) \mathbf{s}^{(g)} + \alpha_u \mathbf{P}^{(g)} \mathbf{D}^{(g)} \mathbf{z}_{\text{isel}}^{(g)}$$

$$\mathbf{s}_\rho^{(g+1)} = (1 - \alpha_s) \mathbf{s}_\rho^{(g)} + \alpha_u \mathbf{B}^{(g)} \mathbf{z}_{\text{isel}}^{(g)}$$

$$\mathbf{H}^{(g+1)} = (1 - \alpha_C) \mathbf{H}^{(g)} + \alpha_C \mathbf{S}^{(g+1)}$$

$$\mathbf{S}^{(g+1)} = \mathbf{s} \mathbf{s}^T$$

$$\rho^{(g+1)} = \rho^{(g)} e^{\gamma \left(\|\mathbf{s}_\rho^{(g+1)}\| - \bar{\rho} \right)}$$



σ correlato tra i vari cromosomi: parametri strategici



- Mutazioni correlate; la correlazione viene rappresentata nella matrice di covarianza.
 - Cromosomi: [x_1, \dots, x_n ; $\sigma_1, \dots, \sigma_n$; $\alpha_{12}, \alpha_{13}, \dots, \alpha_{N-1, N-1}$].
 - La matrice di covarianza $\mathbf{C}_{n \times n}$ gode di queste proprietà:
 - simmetrica
 - $c_{ii} = \sigma_i^2$
 - $c_{ij} = 0$ se i e j sono scorrelati.
 - $c_{ij} \neq 0$, $i \neq j$, c_{ij} viene codificato dai parametri α_{ij}^2 .
 - In questa prima fase supporremo i parametri $\alpha_{ij}^2 = 0$.



Mutazione



$$\mathbf{x}_k^{(g+1)} = \mathbf{x}_j^{(g)} + \rho^{(g)} \mathbf{B} \mathbf{D} \mathbf{z}_k$$

$$\mathbf{z}_k = [z_1 = N(0,1), z_2 = N(0,1), \dots, z_n = N(0,1)]^T \quad \text{Gaussiana}$$

- Vengono generati λ figli, \mathbf{x}_k , a partire dalla mutazione di un genitore, \mathbf{x}_j .
- $\mathbf{x}^{(g+1)}$ sarà funzione dei padri selezionati
- $\rho^{(g)}$ è globale: global step size
- \mathbf{B} \mathbf{D} definiscono l'iper-ellissoide all'interno del quale effettuare le mutazioni.

Devo determinare ancora ρ . Per questo utilizzo l'"evolution path"



Adattamento del global step size ρ



$\sigma_\rho^{(g+1)}$ Evolution path normalizzato (contiene solamente informazioni sulla direzione).
Considero solo B e non D.

$$\mathbf{s}_\rho^{(g+1)} = (1 - \alpha_s) \mathbf{s}_\rho^{(g)} + \alpha_s \mathbf{B}^{(g)} \mathbf{z}_{i_{sel}}^{(g)}$$

$$\bar{e}_n = \left(1 - \frac{1}{4n} + \frac{1}{21n^2} \right)$$

- Approssimazione al 2° ordine del valore atteso della distribuzione della lunghezza dei vettori estratti da $N(\mathbf{0}, \mathbf{I})$

$$\rho^{(g+1)} = \rho^{(g)} e^{\gamma \left(\|\mathbf{s}_\rho^{(g+1)}\| - \bar{e} \right)}$$

$\gamma = 0.1$

- ρ decresce quando la direzione dell'evolution path cambia frequentemente. Quando una stessa direzione è ripetutamente selezionata ρ è incrementato.

Se l'evoluzione cambia spesso direzione, meglio generare figli più vicini (e saranno sparsi lungo tutte le direzioni).



Aggiornamento di α

Regola l'aggiornamento di ρ . Questo deve essere particolarmente sensibile all'innovazione, in modo da pilotare le generazioni future dove più serve.

$$\mathbf{s}_\rho^{(g+1)} = (1 - \alpha_s) \mathbf{s}_\rho^{(g)} + \alpha_u \mathbf{B}^{(g)} \mathbf{z}_{\text{isel}}^{(g)}$$

$$a_u^2 + (1 - a_s)^2 = 1$$

$$a_s = 1 / \sqrt{n}$$

Dove n è la dimensionalità dello spazio della popolazione

$$\alpha_c = 1 / n^2$$

Il numero di parametri liberi della matrice di covarianza è dell'ordine di $n^2 \rightarrow 1/n^2$.

Rappresentano due fenomeni diversi: s è sensibile alle variazioni rapide della fitness della popolazione, mentre h è più sensibile alle variazioni lente, filtrando perciò le variazioni brusche occasionali.

- Es: $n=10$
 - $\alpha_c=0.01, \alpha_s=1/\sqrt{10}=0.316 \rightarrow a_u=0.72$
- La matrice di covarianza ha una *time span* molto più ampia dell'evolution path $O(n^2)$
- L'evolution path può subire cambiamenti veloci mentre la matrice di covarianza ha bisogno di molto più tempo per fissare la memoria della direzione (deve essere selezionata molte volte)

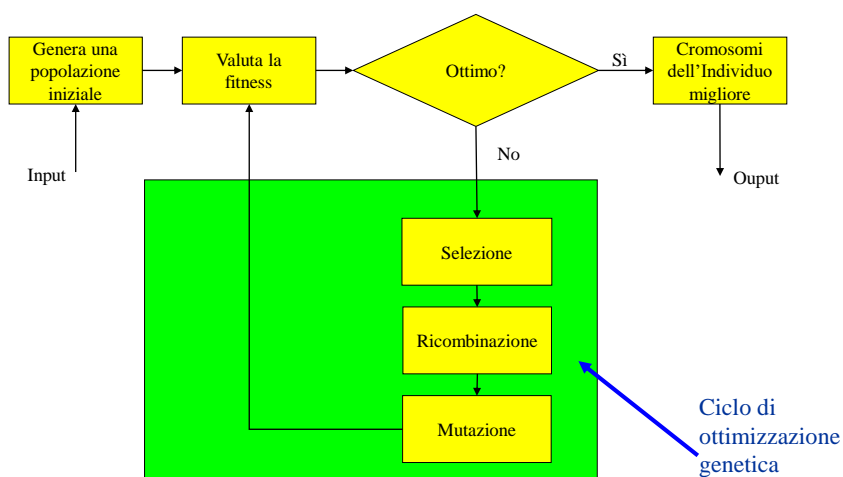
A.A. 2023-2024

67/76

<http://borghese.di.unimi.it/>



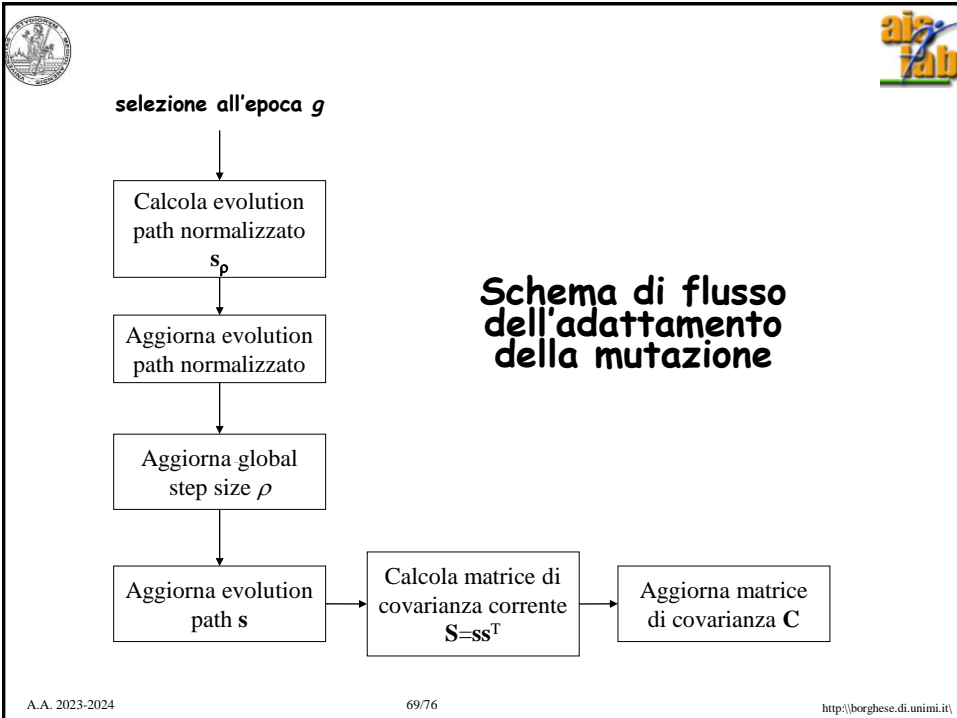
Ottimizzazione evolutiva - flow chart



A.A. 2023-2024

68/76

<http://borghese.di.unimi.it/>





Problema 1: Calibrazione geometrica di una coppia stereo di telecamere



- Determinazione dei parametri geometrici del set-up:
 - Posizione e orientamento relativi di una coppia di camera.
 - Lunghezza focale e punto principale di ciascuna camera.
- Utilizzando la geometria epolare
 - Relazione tra punti omologhi sulle due telecamere
 - Matrice fondamentale 3×3 (7 parametri liberi)
 - 3 rotazioni, 3 traslazioni a meno di un fattore di scala (2Dofs), 2 lunghezze focali
 - **Consente di calcolare 7 parametri in forma chiusa utilizzando la geometria proiettiva.**
- Punti principali
 - Stimo a posteriori i 2 punti principali, c_1 e c_2 , tramite **Ricerca evolutiva**

Spazio della ricerca quadridimensionale, $c_1[x_1, y_1]$ e $c_2[x_2, y_2]$ ("mondo" della popolazione).

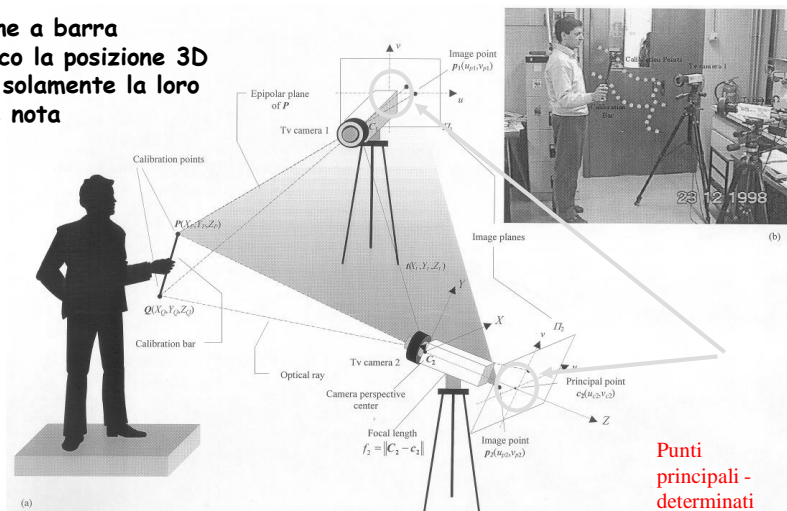
More information in: Cerveri, Pedotti, Borghese, IEEE Trans. Evol. Comp. 2001.




Il set-up




Calibrazione a barra
 Non conosco la posizione 3D dei punti, solamente la loro distanza è nota



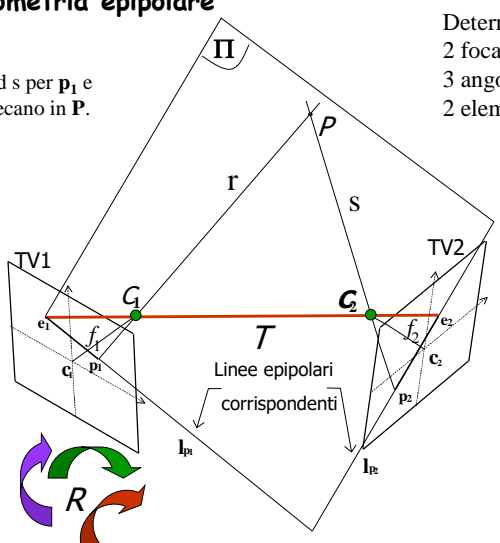
Punti principali - determinati mediante ES



Geometria epipolare



Le rette r ed s per \mathbf{p}_1 e \mathbf{p}_2 si intersecano in \mathbf{P} .




Determino 7 parametri:
 2 focali
 3 angoli di orientamento
 2 elementi del vettore posizione relativa

I parametri sono contenuti nella matrice F


$$\overline{PC_2} \cdot \overline{C_1C_2} \wedge \overline{PC_1} = \mathbf{p}_2^T \mathbf{K}_2^{-T} \mathbf{R}(\mathbf{T} \wedge) \mathbf{K}_1^{-1} \mathbf{p}_1 = 0 \Rightarrow \mathbf{p}_2^T \mathbf{F} \mathbf{p}_1 = 0$$

p_1 e p_2 sono misurati rispetto a c_1 e c_2 non noti.

A.A. 2023-2024
73/76
http://borgnese.di.unimi.it/



Soluzione tramite "Evolution strategies"



Procedura: ES(1,10)

1. scelgo una possibile soluzione per la coppia di punti principali
2. calcolo 7 parametri geometrici in forma chiusa, tramite matrice fondamentale.
3. effettuo la ricostruzione 3D e calcolo il fattore di scala (1 parametro).
4. Valuto la fitness della soluzione.

Funzione fitness: $J = \alpha \sqrt{\frac{1}{N} \sum_{j=1}^N (d_j - L)^2} + \beta \sqrt{\frac{1}{N} \sum_{j=1}^N \left(\frac{e_{int\ j_P}^2 + e_{int\ j_Q}^2}{2} \right)}$

Errore sulla lunghezza della barra.

Errore di intersezione (tra r ed s)

$$\min_{S_2 = \{c_1, c_2\}} J(\{c_1, c_2\} | \{p_1\}, \{p_2\})$$

Elemento della popolazione: $\left\{ \mathbf{w} \left[\mathbf{c}_1(u_{c_1}, v_{c_1}), \mathbf{c}_2(u_{c_2}, v_{c_2}) \right]^T \right\}$

A.A. 2023-2024
74/76
http://borgnese.di.unimi.it/

