

*Progetto*  
*“Casa intelligente per una longevità attiva ed indipendente dell'anziano”*  
*DGR 1464, 7/11/2011*



**Ambient-Aware LiFeStylE tutor, Aiming at a BETter Health**

(Tutoraggio dello stile di vita basato sulla intelligenza ambientale, per una salute migliore)

Risultato D5.5

**Cucina intelligente: moduli di visione e sensori**

*Rev. 1.1, 20 agosto 2014*



## Introduzione

Il progetto AALISABETH dedica particolare rilievo al monitoraggio delle abitudini alimentari: una corretta alimentazione è la base di uno stile di vita salubre. Per questo motivo la cucina rappresenta uno degli ambienti maggiormente monitorati dai sensori del Sistema AALISABETH.

Gli smart-object specifici per l'ambiente cucina (illustrati in altri documenti) sono:

- MIDA fisso; sistema di pesatura e riconoscimento pietanza
- MIDA mobile; “tovaglietta” rigida per l'identificazione del piatto e dell'utente
- CARDEAfrigoBox; multi-sensore di monitoraggio del frigorifero; informa il Sistema sulla temperatura e l'umidità all'interno del frigo, inoltre registra la presenza di luce, dalla quale si può inferire quando il frigo viene aperto; corredato dalla funzione di identificazione (abbinata all'utilizzo del sensore CARDEAmusa) funge anche da dispositivo per la localizzazione indoor
- CARDEAhobBox; multi-sensore di monitoraggio del piano cottura; registra dati relativi a temperatura e umidità e li trasmette al DataBase del Sistema; il sistema prevede una doppia coppia di sensori temperatura-umidità: una coppia serve per la misura dei parametri ambientali e l'altra coppia (dislocata mediante un sondino) per la misura dei parametri in prossimità della cappa
- Sensore magnetico di apertura anta/cassetto

In questo documento si intende documentare l'attività relativa a due ulteriori smart-object/funzioni:

- **SmartCookTop**, ovvero il piano cottura intelligente, sviluppato da ERS
- **MAFI**, ovvero “Monitoring and Analyzing Food Intake”, un'applicazione software finalizzata alla stima automatica della quantità di cibo contenuta in un piatto, sviluppata da UNIPR con la collaborazione di METEDA

## SmartCookTop

Questo smart-object consiste in una piastra fornelli a 4 fuochi. Ogni fornello è munito di una termocoppia per la gestione del controllo dell'accensione dei fuochi, come già avviene nei prodotti commerciali standard. Tuttavia, questo piano cottura ha un sistema di monitoraggio elettronico che può leggere le informazioni relative allo stato delle termocoppie e può comunicarle ad un sistema di

controllo remoto. Si noti che il sistema è unidirezionale, ovvero le informazioni possono fluire dal piano cottura all'esterno, ma il piano cottura non può essere controllato mediante comandi esterni.

La funzione di monitoraggio comprende:

- accensione fuochi (on-off)
- posizione manopola (4 posizioni corrispondenti a fuoco spento, fuoco basso, fuoco medio, fuoco massimo)
- pressione manopola per accensione fornello
- temperatura dei fornelli (misurata mediante la termocoppia integrata in ciascun fuoco).

Lo SmartCookTop dialoga verso l'esterno mediante un collegamento seriale RS232. Tuttavia, per essere pienamente integrato nel Sistema AALISABETH, si è dovuto completare lo smart-object con una periferica di rice-trasmissione wireless in tecnologia ZigBee, fornito da UNIPR.

Il modulo Zigbee è munito di una porta RS232, dialoga con il sistema di monitoraggio fuochi e dialoga con il DataBase centrale per fornire le informazioni al Sistema.

Di seguito, il protocollo di comunicazione per la connessione con il modulo ZigBee.

### Communication protocol

Ref.: Diego Cerioni

rev	date	description	author
2	10/17/14	first emission	VB/EO

### phisycal

Type	Voltage (Vdc)	Baude Rate	No. bit x Byte	Parity	Stop bit	Bit 0 / 1	Start / Stop
UART	3V3	9600	10 (8,1,1)	N	1	0 / 3V3	0 / 3V3

La comunicazione viene sempre avviata dal modulo Zigbee (master)	Il cooktop risponde quando interrogato (slave)
--	--

Command	Read Req	Read Res
Value	0xA0	0X0A

Read Command Request
----------------------

byte index	byte name		byte value	bit index	value description
0	start byte		0x55	na	0x55 = frame start
1	Command byte		0xA0	na	0xA0 = command read req
2	checksum		0-255	na	sum of preceeding bytes

Read Command Response					
0	start byte		0x55	na	0x55 = frame start
1	Command byte		0x0A	na	0x0A = command read res
2	hob1 status		0x00	na	off
			0x01		on
3	hob1 thermocouple		0-255	na	thermocouple adc
4	knob1 status		0x00	na	unpressed
			0x01		pressed
5	knob1 position		0-255	na	position
6	hob2 status		0x00	na	off
			0x01		on
7	hob2 thermocouple		0-255	na	thermocouple adc
8	knob2 status		0x00	na	unpressed
			0x01		pressed
9	knob2 position		0-255	na	position
10	hob3 status		0x00	na	off
			0x01		on
11	hob3 thermocouple		0-255	na	thermocouple adc
12	knob3 status		0x00	na	unpressed
			0x01		pressed
13	knob3 position		0-255	na	position
14	hob4 status		0x00	na	off
			0x01		on
15	hob4 thermocouple		0-255	na	thermocouple adc
16	knob4 status		0x00	na	unpressed
			0x01		pressed
17	knob4 position		0-255	na	position
18	checksum		0-255	na	sum of preceeding bytes

## Monitoring and Analyzing Food Intake

Il contesto di questa applicazione è quello di trovare un modo per aiutare le persone a monitorare e analizzare l'assunzione di cibo. L'idea di base è quella di utilizzare le immagini, prese dall'utente stesso tramite uno smartphone, e di dedurre informazioni qualitative e quantitative circa la pietanza contenuta nel piatto.

Diversi gruppi di ricerca e aziende nel mondo stanno lavorando ad applicazioni simili e questo dimostra quanto il tema sia da considerarsi importante e assolutamente centrale nel contesto dei sistemi di tutoraggio di supporto all'adozione di corretti stili alimentari.

Anche se i diversi gruppi seguono approcci e metodi differenti le sfide da vincere sono quasi sempre le stesse. Il tema può essere suddiviso in tre fasi principali:

1. La segmentazione, ovvero l'identificazione dell'elemento cibo all'interno del piatto e, più in generale, all'interno della foto
2. La classificazione, ovvero il riconoscimento del tipo di cibo
3. La stima del peso del cibo.

Nel progetto AALISABETH si assume come ipotesi che l'utente svolga la funzione di classificazione, comunicando allo smartphone il tipo di pietanza che sta per assumere. Quindi, l'applicazione da sviluppare si dovrà occupare dei punti 1 e 3.

## Segmentazione

Distinguere, in un'immagine fotografica, il cibo da ciò che non lo è, è un compito molto impegnativo. Le immagini possono essere scattate con diverse condizioni di illuminazione, in ambienti differenti, con piatti e toglie molto diverse, ecc...

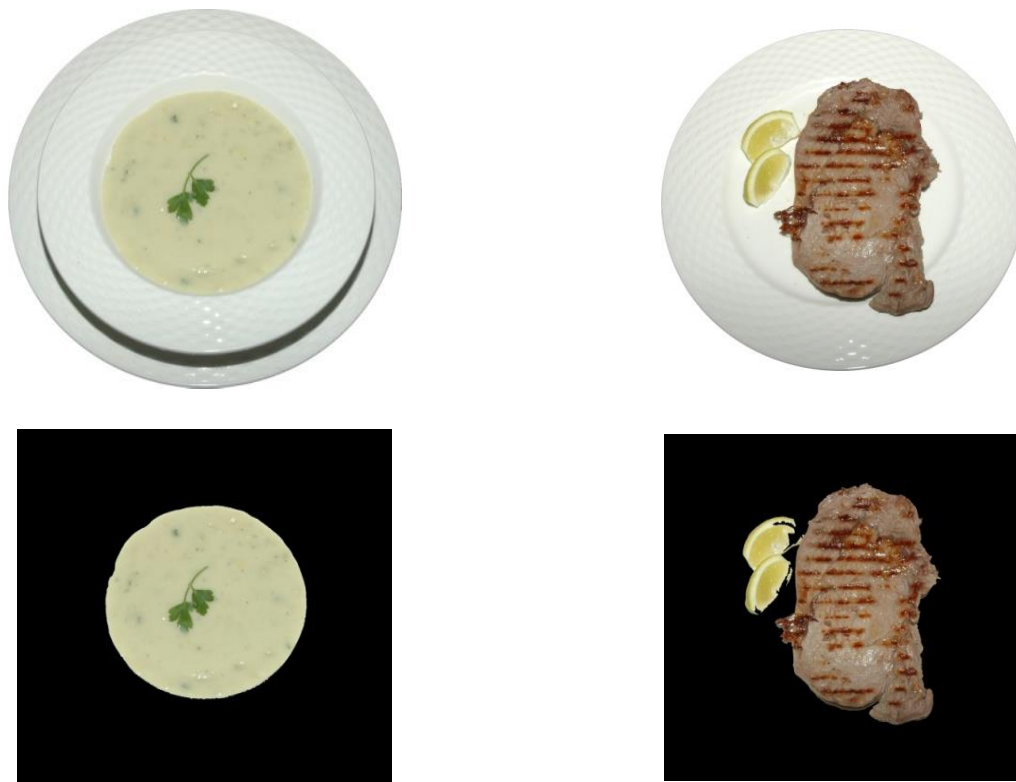
Questi fattori hanno un grandissimo impatto su tutti gli algoritmi di visione artificiale ed in particolar modo sulla funzione di segmentazione.

Si sono presi in considerazione due diversi approcci:

1. Introdurre dei vincoli di variabilità, ovvero fornire all'utente dei limiti da rispettare. Costringere gli utenti a prendere le immagini seguendo alcune regole (che regolino, ad esempio, il tipo e il colore dei piatti, le condizioni di luce, la disposizione sul tavolo, ecc...)
2. Dare agli utenti maggiore libertà e invece chiedere loro di partecipare al processo di segmentazione.

Per il primo approccio si è utilizzato un DataBase di cibi italiani, fornito da MeTeDa Srl. Tutte le immagini del database sono state scattate sostanzialmente nelle medesime condizioni. In questa fase si sono utilizzate diverse tecniche algoritmiche, tra cui K-Means e il filtraggio basato su saturazione

del colore. Si è ottenuta una segmentazione molto precisa in più del 95% delle immagini. Tuttavia le immagini sono prese in uno stesso stato e le caratteristiche generali sono piuttosto simili.

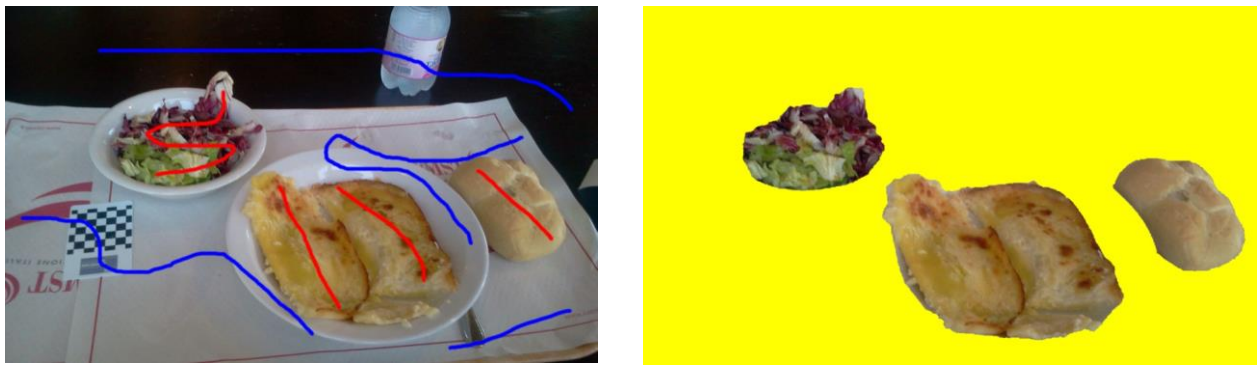


**Figura 1: Segmentazione di immagini METeDa - Le immagini hanno le stesse caratteristiche.**

Per il secondo approccio, si è sperimentata la segmentazione basata su informazioni fornite dall'utente direttamente sull'immagine.

L'idea è quella di chiedere all'utente di marcare, direttamente sul suo smartphone al momento in cui scatta la foto, alcune parti di prodotti alimentari e alcune parti dello sfondo.

Il risultato sulle immagini campione è stato soddisfacente. Si è ottenuta una segmentazione accettabile nella maggior parte delle immagini. Tuttavia, al momento in cui questo documento è stato scritto, non si è ancora avuto modo di sperimentare l'algoritmo su un insieme più grande di immagini reali.



**Figura 2: Segmentazione di immagini pre-elaborate dall'utente (indicazioni blu e rosse)**

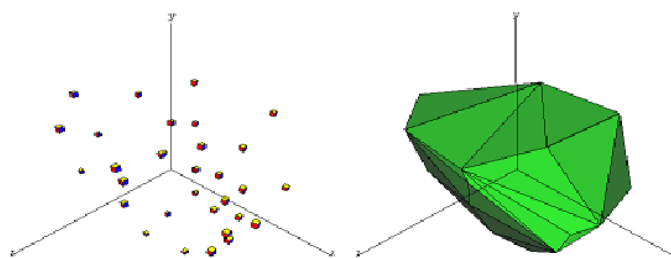
Per calcolare la quantità di assunzione di cibo, abbiamo prima bisogno di stimarne il volume. In questo problema abbiamo considerato due diversi approcci:

1. stima sulla base di ricostruzione in 3D del cibo
2. stima basata su una scena 2D.

Per stimare il volume sulla base di una vista 3D, abbiamo utilizzato due immagini prese dalla stessa scena con la stessa telecamera. Innanzitutto abbiamo ottenuto la ricostruzione 3D utilizzando una sequenza di algoritmi di elaborazione delle immagini. Poi abbiamo cercato di calcolare il volume utilizzando i metodi matematici, come Convex Hull e triangolazione di Delaunay.



**Figura 3: procedura di ricostruzione 3D**



**Figura 4: Stima del volume utilizzando l'algoritmo Convex Hull**

In questo approccio la sfida principale è quella di rimuovere il rumore dell'immagine e i punti ricostruiti imprecisi. Aumentando il numero di immagini si ottiene un risultato migliore. Tuttavia, il costo di trasformazione è troppo alto. Nei nostri test, circa il 50% dei risultati non erano accettabili, soprattutto a causa dei rumori.

Il secondo approccio era basato sull'utilizzo di una sola immagine. Esso si basa sul calcolo della superficie dell'alimento e cercando di stimare la reale quantità di cibo utilizzando alcuni dati di contesto. I dati di contesto potrebbe essere il solito modo di servire un alimento specifico o una regressione generale tra un tipo di cibo e la sua superficie.

Al primo passo è stata calcolata la superficie del cibo. Nelle immagini di prova che abbiamo avuto circa il 90% di accuratezza. Tuttavia, questo risultato deve essere integrato con alcuni dati supplementari per ottenere la quantità di cibo.

In un caso di test, si è cercato di valutare un approccio di regressione per calcolare la quantità di cibo con la sua superficie. Lo abbiamo fatto sul database MeTeDa. Il database dispone di 3 immagini diverse per ogni alimento. La stima del volume è stata fatta con una media di errore del 12%, che è un risultato accettabile. Tuttavia, considerando la dimensione del database e la regolarità delle immagini, sarà necessario valutare questo approccio su un database di immagini più realistiche.

### Checker-board Detection

Utilizzando una fotocamera calibrata, la stima della dimensione è possibile solo in senso relativo se non si dispone di una scala di riferimento. Quindi, si è deciso di utilizzare un riferimento di formato per ottenere la dimensione effettiva: una checker-board.



**Figura 5:** Checker-board usata come riferimento di misura.



L'utilizzo di una checker-board è un metodo semplice per rilevare e sembra essere una scelta appropriata come riferimento di misura. Ci sono anche alcune librerie disponibili per il rilevamento di una scacchiera in un'immagine. Pertanto, si è progettata una scheda standard in PVC (formato carta di credito), con una piccola scacchiera.

Tuttavia, gli algoritmi disponibili, incluso quello in OpenCV, non sono riusciti a rilevare più del 50% delle nostre immagini. Questo problema è dovuto alla calibrazione della telecamera e alle dimensioni ridotte della scacchiera.



**Figura 6: Uso di Checker-board per la calibrazione della telecamera.**

Si è dunque progettato un algoritmo basato su tecniche di intelligenza artificiale per trovare la nostra scacchiera nelle immagini. È stato progettato per essere veloce ed indipendente per la dimensione dell'immagine.



**Figura 7: L'algoritmo realizzato riesce a trovare la scacchiera. OpenCV non riusciva a rilevarla.**

In conclusione, nello sviluppo di questa applicazione di visione ci si è concentrati sulla segmentazione di immagini (rilevamento del cibo) e la stima del volume del cibo. Per ognuna delle due attività si è proceduto seguendo un paio di approcci diversi.

Per la segmentazione, sembra essere vantaggioso chiedere all'utente di pre-elaborare l'immagine scattata, per esempio chiedendogli di tracciare con un dito sullo smartphone le zone della foto in cui è necessario cercare il cibo.

Per la stima del volume, c'è bisogno di risolvere alcuni problemi tecnici. Nella ricostruzione 3D, c'è bisogno di trovare un modo per ridurre il rumore di fondo e ottenere risultati più stabili. Nell'approccio 2D, si deve invece trovare un modo affidabile per integrare informazioni di contesto negli algoritmi.

Inoltre, si è sviluppato un algoritmo per il rilevamento di piccole-scacchiere che possono essere utilizzati nella algoritmo di stima del volume finale.

Per il futuro, c'è la necessità di definire e, auspicabilmente, costruire un DataBase di immagini realistiche che copra interamente il dominio di applicazione. Inoltre, si dovrà pensare a implementare una vera e propria APP per smartphone, per valutare che l'approccio identificato sia sostenibile anche su dispositivi mobili, e che non sia troppo oneroso in termini computazionali.