

Projecte fi de grau

Estudi: Grau en Enginyeria Informàtica

Títol: Implementació de tècniques de percepció
NDT i comparativa amb tècniques ICP

Document: Resum

Alumne: Ricard Segura Duran

Tutor: Pau Vial Serrat
Departament: Arquitectura i Tecnologia
de Computadors
Àrea: Arquitectura i Tecnologia de Computadors

Convocatòria (mes/any): Juny 2022

Resum

Introducció

L'entorn submarí és un domini complex i desfavorable per a les aplicacions robòtiques, ja que la majoria de sistemes de percepció i comunicació utilitzats en l'àmbit terrestre no hi funcionen. L'aigua actua com un gran atenuador de les ones electromagnètiques i, en conseqüència, els sistemes de comunicació Wi-Fi o de localització GPS no hi funcionen. A més, si es volen fer servir sensors làsers subaquàtics - que encara es troben en fase de desenvolupament - s'han d'emprar feixos molt potents que són perillosos per a qualsevol persona o animal que hi estigui a prop. Per tant, sota l'aigua cal recórrer a fenòmens acústics per construir sistemes de comunicació i percepció, ja que les ones mecàniques sí que s'hi propaguen bé. No obstant això, les ones mecàniques són ordres de magnitud més lentes que les ones electromagnètiques i, en conseqüència, proporcionen menys dades i molt més sorolloses. En conseqüència, els algorismes que les processen han de ser més avançats, amb una robustesa i complexitat superiors.

Els Vehicles Submarins Autònoms (AUV) com el Sparus II (Figura 1), són una tecnologia a l'alça per a l'exploració d'entorns submarins. Els AUV són plataformes robòtiques dotades de sistemes de percepció i control que operen sota l'aigua i que permeten efectuar diferents tasques com la localització i la planificació d'una trajectòria de manera autònoma. Un cop establerta la missió a realitzar, el robot es submergeix i comença a executar la tasca definida. La intervenció humana es limita a la supervisió, amb una comunicació mínima que només permet conèixer l'estat del robot. Muntant sensors de rang acústics en un AUV podem efectuar tasques de mapeig per construir mapes d'ocupació de l'entorn del robot.

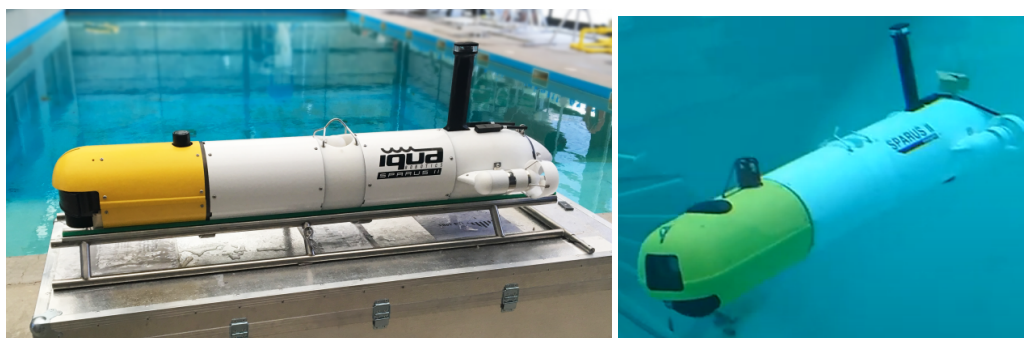


Figura 1: AUV Sparus II desenvolupat al CIRS

Quan es volen generar mapes a partir de l'associació de vistes adquirides (Figura 2) amb un sensor de rang, com pot ser un sonar, si només confiem en el sistema de navegació a la deriva del robot s'obtenen mapes incoherents. Això és degut a que la integració dels sensors inercials del robot genera errors de posició grans. Per evitar-ho, cal dotar el robot d'un sistema d'associació de vistes que permeti determinar el desplaçament del robot quan es registren dues vistes.

L'objectiu principal d'aquest projecte és el desenvolupament, de forma conjunta amb l'estudiant Miguel Malagón, d'una llibreria implementada en llenguatge C++ on s'implementin variants de la tècnica de registre de núvols de punts coneguda com a Normal Distributions Transform per adaptar-la al món acústic. Aquesta implementació ha de ser capaç d'executar-

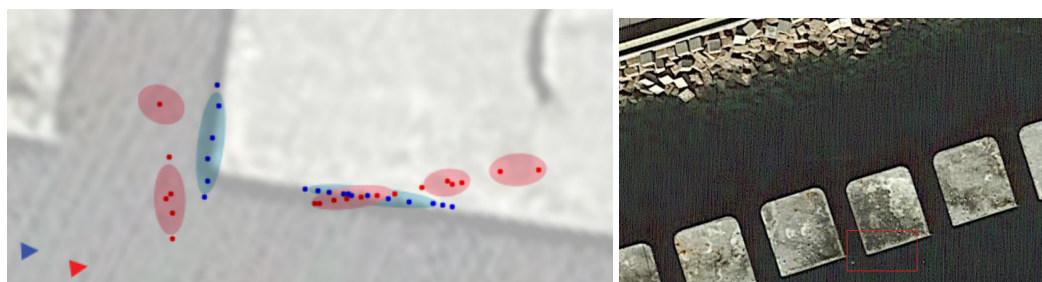


Figura 2: Exemple vistes adquirides en un entorn estructurat (Port de Sant Feliu de Guíxols)

se en temps real per tal de que el robot sigui capaç de generar en línia el mapa volumètric del seu entorn.

Metodologia

Per al desenvolupament del projecte s'ha seguit una metodologia de desenvolupament de tipus àgil, on s'han definit sprints setmanals per realitzar les diferents tasques. S'ha realitzat d'una forma iterativa i incremental, on cada versió nova que es genera millora la versió anterior del projecte; refinant i millorant les diferents parts i on, alhora, s'implementen noves funcionalitats. La introducció a l'àmbit de treball i a la tècnica de registre s'ha fet de manera conjunta amb l'estudiant Miguel Malagón. Això ens ha permès trencar el gel dels coneixements teòrics en companyia. Alhora, s'ha realitzat conjuntament el disseny de l'estructura de la llibreria i la implementació dels objectes bàsics. Aquest fet ens ha permès acordar l'estructura bàsica de la llibreria que després evolucionarem per separat. Definit el punt de partida, aquest projecte s'ha especialitzat en comparar les tècniques implementades en aquesta llibreria amb altres tècniques de l'estat de l'art ja implementades en llibreries públiques no especialitzades en l'entorn submarí. Feta la comparativa, s'ha tancat el projecte introduint nous algorismes procedents de la comunitat de l'aprenentatge automàtic no supervisat.

Disseny i implementació base de la llibreria

A partir d'una primera implementació realitzada seguint les explicacions d'un article científic, es va definir una estructuració de les classes que formen la llibreria. Aquesta estructura defineix tres tipus d'objectes conceptuals, el Front-End, el Mètode de registre i el Solucionador.

El Front-End s'encarrega de fitar un Gaussian Mixture Model (GMM) a un núvol de punts. Un GMM és un model probabilístic format per un sumatori de diverses distribucions gaussianes. És un model continu que permet encapsular la informació discretitzada del núvol de punts.

El Mètode de registre defineix el problema de registre com un problema d'optimització. S'estableix com a variable de decisió la transformació entre els dos escanejos. Es defineixen dues funcions de cost. El mètode Point to Distribution (P2D) registra un escaneig parametritzat com a núvol de punts contra un altre escaneig parametritzat com a GMM, maximitzant el likelihood. En canvi, el mètode Distribution to Distribution (D2D) registra dos scans parametritzats com a GMM, minimitzant la seva divergència.

El Solucionador s'encarrega de resoldre un problema d'optimització aplicant mètodes numèrics. En aquest projecte ens centrem en mètodes del gradient, basats en les derivades de la funció de cost.

A la figura 3 es mostra l'estructura de classes que s'ha definit per la llibreria. La classe base de la llibreria és **GaussianMixturesModel** que és una estructura de dades que permet

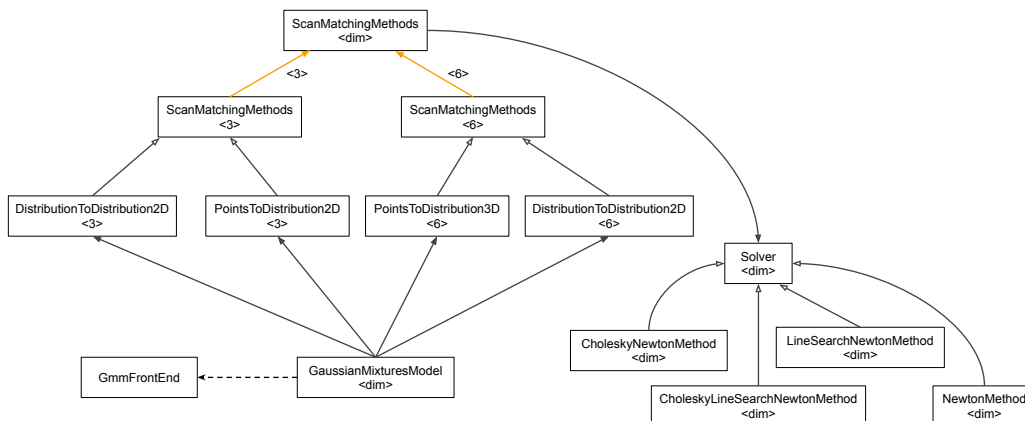


Figura 3: Estructura de classes definida.

emmagatzemar un GMM.

L'objecte `GaussianMixturesModel` conté les dades necessàries que conformen un GMM. Aquestes dades són: K nombre de components del model, $\phi_{i=1\dots K}$ pes de cada component on la suma de totes ha de ser 1, $\mu_{i=1\dots K}$ la mitjana del component i i $\sigma_{i=1\dots K}$ covariància del component i . Aquests paràmetres es guarden com atributs de l'objecte.

Com que no hi ha una única tècnica per fitar un GMM a un núvol de punts, s'ha dissenyat el Front-End com una llibreria de funcions. Totes les funcions reben com a argument un núvol de punts i retornen un objecte de tipus `GaussianMixturesModel`. El Mètode de registre s'ha dissenyat com una interfície, anomenada `ScanMatchingMethod`. Utilitzar la interfície permet al solucionador determinar la funció de cost i les seves derivades independentment del mètode implementat. La interfície es templatitza segons la variable dimensió dels punts de l'escaneig, encara que es faran implementacions separades dels casos bidimensionals i tridimensionals per la impossibilitat de templatitzar el codi de les derivades. El Solucionador s'ha dissenyat com a interfície per poder implementar diferents algorismes. La variable de decisió del problema s'ha templatitzat segons dimensió per poder-lo aplicar indistintivament a problemes bidimensionals o tridimensionals.

De la llibreria Front-End s'ha implementat la funció anomenada `'ndt_constructor()'`. Aquesta funció fita un GMM seguint el planteig de la tècnica NDT. Aquesta tècnica projecta una graella cartesiana sobre el núvol de punts i estableix una component gaussiana a cada casella amb un mínim nombre de punts.

De la interfície `ScanMatchingMethod` s'han implementat dues subclasses templatitzades en dues dimensions: `Point2Distribution2D` i `Distribution2Distribution2D`, que implementen respectivament les tècniques P2D i D2D bi-dimensionals. Gràcies a la interfície ambdues classes proporcionen el mateix. Al constructor es proporcionen els escanejos a registrar amb la parametrització adequada i els mètodes `'compute_score()'`, `'compute_score_and_gradient()'` i `'compute_score_gradient_and_hessian()'` retornen el valor de la funció de cost i de les seves derivades, donada per paràmetre una transformació entre els dos escanejos a registrar.

De la interfície de Solucionador s'han implementat dues subclasses: `NewtonMethod` i `LineSearchNewtonMethod`. `NewtonMethod` implementa el mètode de Newton bàsic, mentre que la subclasse `LineSearchNewtonMethod` inclou al mètode de Newton un algorisme per determinar la mida òptima del pas d'optimització a cada iteració a partir de les condicions de Wolfe. Gràcies a la interfície, totes les subclasses tenen com a atribut una instància

del mètode de registre a resoldre i implementen el mètode `'compute_optimum()'`. Aquest mètode implementa l'algorisme d'optimització triat.

El codi de la llibreria s'ha documentat amb el generador de documentació Doxygen. Aquest programa permet generar un document de tipus HTML o PDF amb la informació proporcionada amb comentaris tipus al codi. Aquesta tasca de documentació és útil ja que quan la llibreria estigui enllestida es vol fer pública perquè pugui ser utilitzada per altres investigadors.

Comparativa amb altres tècniques de registre

Un cop implementada la base de la llibreria, s'ha fet una comparativa entre la implementació proposada amb altres tècniques habituals en l'estat de l'art (figura 4). Aquesta comparativa s'ha dut a terme per validar la qualitat de la proposta realitzada i per veure com s'ajusta cada tècnica a les limitacions del domini submarí. La comparació s'ha realitzat amb les tècniques Iterative Closest Point (ICP) i Generalized ICP (GICP) que són àmpliament aplicades a la comunitat de la robòtica i es troben implementades en la llibreria de codi obert Point Cloud Library.

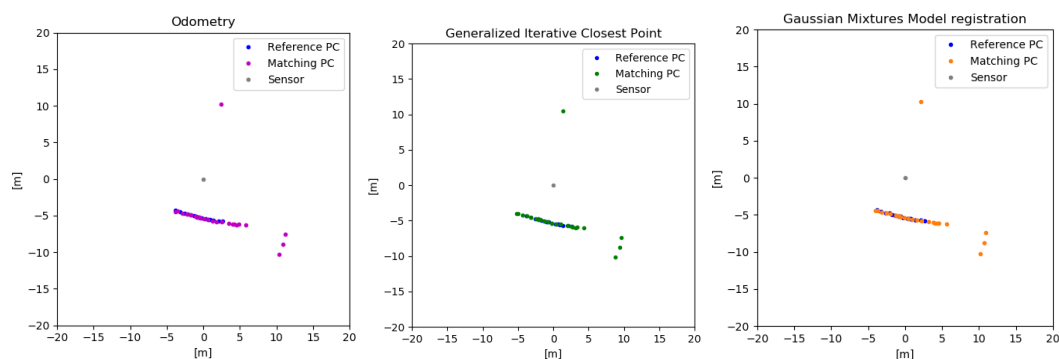


Figura 4: Exemple de resultats obtinguts

Per dur a terme les comparacions entre les diverses tècniques, s'han fet servir data sets amb dades reals adquirides amb un AUV. Com que aquests data sets han estat adquirits amb l'AUV submergit, no es disposa de senyal de GPS, ja que les ones electromagnètiques només penetren pocs centímetres dins de l'aigua. Per la qual cosa, no es disposa del ground truth de la trajectòria seguida pel robot. En conseqüència, no es disposa d'una mesura exacta de la trajectòria del robot i no es pot mesurar l'error de cada tècnica de registre. Llavors, la comparativa només pot ser qualitativa basant-se en l'observació visual dels registres assolits.

Dels resultats obtinguts en fer la comparativa es conclou que la tècnica ICP és la que presenta els resultats més pobres. Falla una gran quantitat de registres en comparació a les altres dues. De les comparatives realitzades entre la tècnica GICP i la implementada en aquest projecte, s'ha vist tant per un entorn estructurat com en un entorn desestructurat les dues tècniques obtenen uns resultats molt bons. En general no hi ha gaire diferència en termes de resultats entre les dues tècniques. Per la qual cosa es conclou que les dues tècniques s'adapten bé, amb la diferència que la tècnica implementada en aquest projecte retorna una mesura de la incertesa del registre. Que podrà ser utilitzada per a millorar la navegació del vehicle.

Ampliació de la llibreria de Front-Ends

En finalitzar la comparativa es va decidir realitzar una ampliació de la llibreria de Front-Ends. Les noves funcions apliquen els algorismes K -means i Expectació- Maximització (EM)

per fitar un GMM a un núvol de punts. L'algorisme K -means té com a objectiu la classificació de conjunts amb N punts en K grups, on cada punt s'assigna només al grup més proper segons la distància euclídia (equació 1).

$$J = \sum_{n=1}^N \sum_{k=1}^K r_{nk} \|x_n - \mu_k\|^2 \quad (1)$$

Aquesta classificació de les dades pot servir per fitar-hi un GMM, establint una component gaussiana per cada grup. L'algorisme d'EM és un mètode que classifica de manera probabilística les associacions dels punts respecte als grups fent associacions no rígides (equació 2), on un punt pertany a tots els grup en una quantitat.

$$\gamma(z_{nk}) = \frac{\pi_k \mathcal{N}(x_n | \mu_k, \Sigma_k)}{\sum_j \pi_j \mathcal{N}(x_n | \mu_j, \Sigma_j)} \quad (2)$$

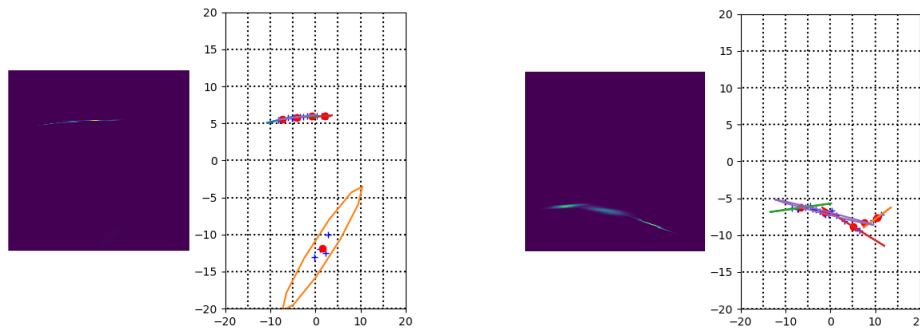


Figura 5: Exemple de GMMs amb K -means i EM

Les dues funcions s'han implementat seguint una mateixa estructura. A l'inici es creen les variables on es guardaran els resultats per als GMM. Un cop creades les variables, s'inicien els bucles on es fan els algorismes amb les respectives etapes d'expectació i de maximització per a cada tècnica. Cada cop que es troba una solució es mira si és millor i en cas de ser-ho es desa.

Un cop feta la implementació base de les tècniques s'ha fet un estudi per a cadascuna on es mira quins criteris de parada es poden definir. En el cas del constructor K -means s'han fet servir la inèrcia i la trace de les matrius com a criteri de parada, en el cas del constructor EM s'ha utilitzat el likelihood com a criteri. Per a comprovar que els resultats obtinguts d'aplicar les noves funcions s'han tornat a fer servir els data sets emprats en la comparativa. Amb els resultats aconseguits s'ha pogut veure com s'han fitat els GMM d'una forma més flexible, pel fet que no s'ha fet servir la graella que generava el constructor NDT. En la figura 5 es poden veure GMMs resultants en aplicar aquesta tècnica.

Conclusions

En conclusió, en aquest treball s'ha realitzat la implementació d'una llibreria en C++ basada en la tècnica NDT. Alhora, s'aporten dades de la seva comparativa amb altres tècniques en l'estat de l'art. Finalment, s'ha ampliat la llibreria de funcions Front-End implementant els algorismes de K -means i d'EM.