

Generarea unui model 3D al scenei folosind sisteme sensor fusion

Tudor - Constantin Ivanovici

Rezumat

Definirea problemei

Generarea unui model 3D al unei scene reprezintă procesul de achiziție, procesare și stocare a informațiilor ce cuprind forma și aspectul corpurilor dintr-un mediu fizic. Sistemele eficiente de reconstrucție 3D reprezintă una din problemele de fundamentale ale domeniului computer vision. Conceptul de sensor fusion reprezintă combinarea a mai multor seturi de date provenite din surse separate astfel încât informația rezultată să aibă un nivel mai mic de incertitudine asupra măsurătorilor decât sursele originale. Astfel, o reconstrucție eficientă și de înaltă calitate poate fi obținută prin procesarea mai multor astfel de seturi de date, pentru o înțelegere cât mai exactă asupra mediului analizat.

Procesul de reconstrucție 3D este extrem de util în contextul aplicațiilor bazate pe computer vision și achiziție de date de la unul sau mai mulți senzori, precum sisteme de navigare pentru roboți mobili, tehnologii asistive sau sisteme de conducere autonomă a vehiculelor. De asemenea, aceste tehnici sunt implementate și în cadrul unor tehnologii emergente precum realitatea virtuală și realitatea augmentată. Necesitatea construcției unui model 3D al lumii cât mai fidel acesteia este cu atât mai important cu cât interfețele cu sistemele de calcul devin din ce în ce mai naturale în folosire pentru utilizatori.

Abordare

Lucrarea își propune analiza tipurilor de metode, algoritmi și structuri de date folosite în cadrul sistemelor de reprezentare a datelor 3D bazate pe sensor fusion. Un al doilea obiectiv îl reprezintă descrierea unei implementări a unui astfel de sistem, bazată pe seturi de date ce pot fi achiziționate cu diverși senzori, din care se derivă informația tridimensională.

Rezultate

Algoritmii explorați în implementare descriu un proces de fuziune incrementală aplicat unui model 3D, bazat pe hărți de adâncime calculate din informațiile achiziționate de un sistem de camere stereo RGB. Procesul de reconstrucție se bazează pe folosirea voxelilor pentru a partiționa spațiul tridimensional și folosirea funcțiilor TSDF (Truncated Signed Distance Function) pentru a reprezenta geometria spațiului. Avantajul folosirii acestor funcții este dat de numărul mic de voxelii necesar pentru reprezentarea unei anumite suprafețe în memorie.

Deși sistemele de voxelii pot fi stocate în mai multe modalități, precum folosirea arborilor sau a tablourilor, structura de date folosită, de tip tabelă de dispersie, facilitează viteza acestui mecanism prin operații eficiente executate în paralel pe un procesor grafic.

Pentru evaluarea implementării a fost folosit setul de date KITTI, ce constituie o secvență de imagini stereo, preluată în mișcare de un sistem de camere poziționat pe un autovehicul.

Concluzii

Reconstrucția de timp real a modelelor 3D din lumea fizică reprezintă unul din pașii fundamentali ai procesului de înțelegere vizuală automată ce stă la baza domeniului computer vision. Soluția explorată în lucrarea de față descrie o metodă eficientă de reconstrucție 3D a scenelor, permițând integrarea facilă de date de la mai mulți senzori.