סקירת ספרות

נבחן את הפוטנציאל של מציאות רבודה (augmented reality או AR) בהגברת המודעות למצב (situational awareness או SA) במהלך מבצעים צבאיים.

סביבות צבאיות הן כאוטיות מטבען, עם מידע מקוטע או לקוי, רעש ולחץ שמעמיס על החושים האנושיים ותרחישים משתנים במהירות. גורמים אלו מסבכים את יכולתו של החייל לעבד ולצפות מרכיבים או אירועים בסביבתו, ובסופו של דבר משפיעים על קבלת ההחלטות ויעילותו המבצעית. מציאות רבודה יכולה להיות פתרון לאתגרים אלה על ידי שילוב מידע דיגיטלי עם הראייה הטבעית של חייל לגבי הסביבה בצורה אינטואיטיבית, לא פולשנית ורלוונטית מבחינה הקשרית שתתמוך בהבנה של הסיטואציה אבל תרגיש טבעית בשילוב שלה עם חושי התפיסה האנושיים.

מודעות למצב, כמושג, היא בסיס של פעולות צבאיות אפקטיביות. זה כולל שלושה שלבים היררכיים: תפיסת אלמנטים בסביבה, הבנת האינטראקציות והמשמעות שלהם והקרנת מצביהם העתידיים. מודעות למצב תלויה ביצירת מודלים מנטליים המסייעים לחיילים להתמקד בנושאים קריטיים ולקבל החלטות בזמן תחת לחץ. עם זאת, השגת מודעות בלחימה מודרנית היא מאתגרת בשל כמות המידע העצומה הזמינה. AR יכולה לפחית את הבעיה על ידי סינון ותעדוף נתונים, ומאפשרת או מכוונת חיילים להתמקד במה שחשוב ביותר בכל סיטואציה.

זה מביא אותנו למושג חשוב אחר, מודעות להקשר (context awareness או CA). סוג המודעות הזה מתייחס ליכולת של המערכת להסתגל באופן דינמי לסביבת החייל. זה כולל מיקום, התמצאות ואלמנטים סמוכים. מערכות AR ממנפות זאת על ידי התאמת המידע המוצג על סמך מיקומו ותנועתו של החייל. יכולת הסתגלות זו מבטיחה שהנתונים המוצגים תמיד רלוונטיים וניתנים לפעולה. קוגניציה מוגברת משלימה את היכולות הללו על ידי טיפול במגבלות של עיבוד מידע אנושי. הוא מתמקד בעיצוב ממשקים המציגים מידע ביעילות, תוך מזעור עומס קוגניטיבי.

הבסיס הטכנולוגי של AR ביישומים צבאיים צריך לכלול מערכות המחברות חיילים, כלי רכב ומרכזי פיקוד, ומייצרים תמונה מבצעית משותפת. מערכות אלו משתמשות בחיישנים, מצלמות, והתקני תקשורת וטכניקות עיבוד כדי לקבל, לשדר ולנתח נתונים. השלב האחרון כולל ייצוג חזותי של נתונים אלה באמצעות AR שישפר את התפיסה של החייל לגבי שדה הקרב.

כמה מערכות התומכות ב-AR כבר הוכיחו את הפוטנציאל לחולל מהפכה בפעולות צבאיות. דוגמאות בולטות מופיעות בהמשך.

כדי להפוך מערכות AR ליעילות, יש לעמוד בדרישות מסוימות. אלה כוללים מודלים חזקים של מידע שהופכים נתונים סביבתיים גולמיים לתובנות ניתנות לפעולה, מנגנוני סינון יעילים להפחתת עומס המידע ויכולות תפעוליות כמו אינטראקציה מבוססת מחוות, מעקב בזמן אמת ותקשורת מאובטחת. מערכת ה-AR האידיאלית צריכה להיות גם קלת משקל, ניידת ומסוגלת לתפקד באופן עצמאי תוך הסתמכות מינימלית על רשתות חיצוניות.

לסיכום חלק זה, הפוטנציאל של AR לשפר את המודעות הסביבתית וקבלת החלטות בהקשרים צבאיים הוא עצום, כדאי לניצול ויכול להציל חיי אדם ולהוביל למבצעים צבאיים מדוייקים יותר ברמת המיקרו. על ידי שילוב חלק של מידע דיגיטלי ומידע מהעולם האמיתי, AR מפחיתה עומס קוגניטיבי ומעצים חיילים לקבל החלטות מושכלות, אפילו בסביבות מתח גבוהות. מחקר ופיתוח עתידיים יתמקדו בחידוד טכנולוגיות אלו ליישום רחב יותר, הבטחת יכולת הסתגלותן, חוסנן ושימושיותן על פני תרחישי לחימה מגוונים.

כעת נעבור לבחון טכנולוגיות קיימות ופרוייקטים open source שניתן להתבסס עליהם כשנעבוד על מערכות AR וניתוח וידאו.

אלגוריתמים מתקדמים, בשילוב עם חומרה חזקה, הם חיוניים בהבטחת דיוק וביצועים בהגדרות ברמה אופטימלית.

אלגוריתם YOLO (You Only Look Once):

זהו אלגוריתם זיהוי אובייקטים בזמן אמת המעבד תמונות שבזכות היותו open source מתעדכן כל פעם לטכנולוגיות המתקדמות ביותר. הוא ידוע במהירות ובדיוק שלו ומזהה אובייקטים, קובע את מיקומם בתוך מסגרת ומסווג אותם. היעילות שלו הופכת אותו למתאים למערכות AR שבהן חיילים דורשים עדכונים בזמן אמת. גרסאות YOLO מתקדמות, כגון YOLOv5 או YOLOv8, מציעות דיוק משופר והסקת הסקה מהירה יותר, אידיאלית לזיהוי כלי נשק, כלי רכב ואנשי צוות בשדה הקרב והוא ניתן למימוש גם בפייתון וגם בC, תלוי בצורך.

R-CNN Mask(Region-based Convolutional Neural Network):

Mask R-CNN מרחיב את זיהוי האובייקטים המסורתי על ידי הוספת יכולות סגמנטציה, המאפשרות לה לתחום אובייקטים ברמת פיקסלים. זה יכול לשפר את מודל ה-AR שלנו על ידי זיהוי מדויק של הקצוות של אובייקטים, כגון הבחנה בין מדים לשטח שמסביב. זה שימושי במיוחד בסביבות צפופות או מורכבות כמו לוחמה עירונית.

זרימה אופטית (Optical flow):

טכניקות זרימה אופטית מנתחות את התנועה של אובייקטים בין פריימים רצופים של וידאו. שיטה זו מסייעת לעקוב אחר מטרות נעות, לחזות את מסלולן ולזהות דפוסים בתנועתן. זרימה אופטית חשובה עבור מערכות AR שחייבות לספק לחיילים מידע על איומים דינמיים.

התאמת תכונות:

טכניקה המשמשת בראייה ממוחשבת כדי לזהות ולעקוב אחר מאפיינים על פני מסגרות וידאו עוקבות או בין מספר תמונות. תכונות אלו עשויות לכלול קצוות, פינות או טקסטורות שנשארות עקביות ללא קשר לשינויים בקנה מידה, סיבוב או תאורה. זה הופך את התאמת התכונות לקריטית עבור יישומים שמטרתם היא זיהוי אובייקטים בדיוק הגבוה ביותר האפשרי. אלגוריתמים מרכזיים בהתאמת תכונות כוללים את הבאים:

SIFT (Scale-Invariant Feature Transform):

מזהה נקודות מפתח בתמונה שאינן משתנות לקנה מידה, סיבוב ושינויים קלים בתאורה. האלגוריתם מאתר נקודות עניין באמצעות גישה של Difference of Gaussians לזיהוי אזורים בעלי ניגודיות גבוהה (הבדלים גאוסיים (פונקציות מתמטיות המשמשות ליצירת עקומות חלקות בצורת פעמון) היא טכניקה המשמשת לזהות קצוות ונקודות מפתח על ידי הפחתת שתי גרסאות מטושטשות של אותה תמונה, כל אחת נוצרה עם מסננים גאוסיים של סטיות תקן או קנה מידה שונים. זה מדגיש אזורים עם שינויים מהירים בעוצמה, מה שהופך אותו ליעיל לזיהוי תכונות כמו פינות וקצוות). כלומר תיאורים מתמונות שונות מושווים כדי למצוא התאמות על סמך דמיון.

יצירת מתאר: עבור כל נקודת מפתח, SIFT יוצר וקטור הלוכד שיפועים מקומיים של תמונה סביב הנקודה.

יתרונות: דיוק גבוה וחוסן בפני טרנספורמציות.

חסרונות: אינטנסיבי מבחינה חישובית, יכול להפוך אותו לפחות מתאים ליישומים בזמן אמת.

SURF (Speeded-Up Robust Features):

זו בעצם גרסה אופטימלית של SIFT שנועדה להיות מהירה יותר תוך שמירה על חוסן.

זיהוי נקודות מפתח: SURF משתמש בייצוג תמונה אינטגרלי ובמטריצה ההסיאנית כדי לזהות במהירות תחומי עניין. קווי המתאר מבוססים על התפלגות שינויי העוצמה הסמוכים לנקודות מפתח.

יתרונות: מהיר יותר מ-SIFT ועדיין מציע דיוק טוב.

חסרונות: מעט פחות חזק מ-SIFT בתרחישים מאתגרים מסוימים.

ORB (Oriented FAST and Rotated BRIEF):

ORB הוא שילוב של שתי שיטות יעילות: FAST (Features from Accelerated Segment Test) ו-BRIEF (Binary Robust Independent Elementary Features).

ORB משתמש באלגוריתם FAST כדי לזהות במהירות נקודות מפתח ולהתאים אותן באוריינטציה סיבובית לצורך מודל תלת מימדי ומשתמש ב-BRIEF כדי לייצר מתארים על ידי השוואת עוצמות פיקסלים סביב נקודות מפתח, מה שהופך אותם לקלים ומהירים לחישוב.

יתרונות: מהיר במיוחד ומתאים ליישומים בזמן אמת, כמו מערכות AR בסביבות צבאיות.

חסרונות: עשויים להיות פחות מדויקים מ-SIFT או SURF בזיהוי תכונות עדינות.

ניתן להשתמש באלגוריתמים של התאמת תכונות כדי לעקוב אחרי אובייקטים בדיוק מירבי, זיהוי ציוני דרך, הטמעת רכיבי AR בראייה מרחבית ועוד. אלגוריתמים של התאמת תכונות דורשים תמיכה חישובית, במיוחד כאשר עובדים עם תמונות ברזולוציה גבוהה או הזנות וידאו בזמן אמת. GPUs ומאיצים מיוחדים כמו TPUs יכולים לשפר משמעותית את מהירות העיבוד, ולהבטיח שאלגוריתמים אלה פועלים בצורה חלקה במערכות AR צבאיות.

DeepSORT (Simple Online and Realtime Tracking):

אלגוריתם זה מצוות לעתים קרובות עם YOLO למעקב מרובה אובייקטים. DeepSORT משתמשת בשילוב של סימני תנועה ומראה כדי לשמור על מזהי מעקב עקביים עבור אובייקטים בזמן שהם נעים בסצנה, דבר שיכול להיות קריטי בשמירה על מודעות למצב באזורי לחימה דינמיים.

ניתוח הווידאו של החומרה למערכות AR צבאיות חייב לאזן בין ביצועים, ניידות ויעילות אנרגטית. מרכיבי המפתח יכולים לכלול:

GPUs (יחידות עיבוד גרפיות): GPUs מודרניים, כמו סדרת Jetson של NVIDIA, מותאמים לעיבוד מקביל ואידיאליים להפעלת אלגוריתמי למידה עמוקה כמו YOLO או Mask R-CNN. הגודל שלהם הופך אותם למתאימים למכשירים לבישים.

TPUs: פותחו על ידי Google, מספקים האצה יעילה לחישובי רשת נוירונית וניתן לשלב אותם במערכות AR כדי לשפר את הביצועים בזמן אמת.

מצלמות ברזולוציה גבוהה וראיית לילה או מצלמות תרמיות לשפר ראות בתנאי תאורה או מזג אוויר מעורפלים.

מכשירי Edge כמו NVIDIA Jetson Nano או Xavier מאפשרים לבצע חישובים באופן מקומי במקום להסתמך על עיבוד ענן. זה מקטין את זמן האחזור ומבטיח שחיילים יכולים לגשת למידע קריטי גם באזורים שנפגעו ברשת.

אחסון מהיר (NVMe SSDs) וזיכרון RAM גדול נחוצים כדי להתמודד עם נפח הנתונים הגדול שנוצר על ידי ניתוח וידאו בזמן אמת.

לסיכום, מערכת AR חזקה ליישומים צבאיים דורשת שילוב הדוק ומסונכרן היטב של אלגוריתמים מודרניים לניתוח וידאו וחומרה אופטימלית. אלגוריתמים כמו YOLO ו-Mask R-CNN מספקים זיהוי ופילוח אובייקטים בזמן אמת, בעוד שחומרה כמו GPUs והתקני EDGE מבטיחים פעולה חלקה בסביבות תובעניות. על ידי שילוב טכנולוגיות אלה עם נתונים מהשטח וטיפול באילוצי מערכת, AR יכול לספק לחיילים מודעות מצב ויתרונות טקטיים חסרי תחליף בשדה הקרב, שאדם לא משנה כמה ניסיון יש לו לא יוכל להחליף.

דוגמאות לפרויקטים דומים שקיימים בשוק ועונים על צרכים דומים (סקירת שוק):

Microsoft's Integrated Visual Augmentation System (IVAS):

IVASפותחה עבור צבא ארה"ב. אלו אוזניות מציאות רבודה (augmented reality - מציאות רבודה מכסה מידע או אובייקטים דיגיטליים על העולם האמיתי, ומשפרת את תפיסת המשתמש את סביבתו) שנועדו לשפר את המודעות למצב על ידי הנחת תמונות חיישנים ומידע אחר על שדה הראייה של החייל. הוא משלב את טכנולוגיית HoloLens של מיקרוסופט עם תכונות נוספות המותאמות לשימוש צבאי. המערכת כרגע עוד בשלבי פיתוח ועוברת בדיקות, כאשר השטח צפוי ב-2025.

Vegvisir's Mixed Reality Situational Awareness System

Anduril Industries, חברת טכנולוגיה ביטחונית אמריקאית, משתפת פעולה כדי לשפר אוזניות מציאות רבודה צבאיות על ידי שילוב תוכנת Lattice (מערכת שהם פיתחו המאפשרת אינטגרציה מערכתית של מערכות ושירותים קיימים כדי לאפשר העברת נתונים). Lattice ממזגת נתונים חיים ממערכות הגנה שונות, ומספק מידע בזמן אמת, מודע להקשר לחיילים, ובכך משפר את מודעות המצב ויכולות קבלת ההחלטות שלהם.

Vegvisir's Mixed Reality Situational Awareness System

הפעם דוגמא מאירופה, הסטארט-אפ האסטוני-קרואטי Vegvisir מפתח מערכת מציאות מעורבת (מציאות מעורבת ממזגת את העולמות הפיזיים והדיגיטליים על ידי שילוב סביבות אמיתיות ווירטואליות, ומאפשרת לאובייקטים פיזיים ווירטואליים לקיים אינטראקציה בזמן אמת) שמטרתה להגביר את המודעות למצב עבור צוותי רכב משוריין. המערכת מספקת סקירה של 360 מעלות של סביבת הרכב באמצעות מצלמות מותקנות ברכב ואוזניות אטימות נמוכות במיוחד, המאפשרות התמצאות וקבלת החלטות טובות יותר בשדה הקרב.

Eyekon

מקורו של Eyekon כיוזמה של צבא ארה"ב לשיפור קבלת ההחלטות בתרחישים בשדה הקרב באמצעות AR. המערכת משלבת מחשבים לבישים המנתחים ומציגים מידע בזמן אמת. חיילים משתמשים במערכת Eyekon כדי להמחיש נתונים הקשורים למטרה ומידע קריטי אחר בשדה הקרב ישירות על תצוגות הנשק שלהם. המטרה הטקטית היא לספק עדכונים בזמן אמת, מודעים להקשר, על הסביבה הקרובה של החייל, ולאפשר תגובה מהירה ומדויקת יותר לאיומים. המערכת משתמשת בחיישנים כגון GPS, יחידות מדידה אינרציאליות וגלאי אינפרא אדום כדי לאסוף נתונים.

BARS (Battlefield Augmented Reality System):

פיתוח של מעבדת המחקר הימי של ארה"ב. BARS מתמקדת במבצעים צבאיים בשטח עירוני כאשר המודעות למצב מאתגרת במיוחד. מטרתו העיקרית היא לספק ללוחמים הבנה ברורה של הסביבה שלהם, כולל אלמנטים חסומים או נסתרים. המערכת משתמשת בתצוגות מותקנות בראש כדי להצמיד הערות גרפיות על תצוגות עולם אמיתיות, תוך יישור מידע וירטואלי עם אובייקטים פיזיים בשדה הראייה של המשתמש.

iARM (Intelligent Augmented Reality Model):

מערכת זו פותחה תחת DARPA (defense advanced research projects agency, עוד מתוצרת ארה"ב) ונועדה לשפר את קבלת ההחלטות בסביבות מורכבות ומעוררות תחרות באמצעות מערכת AR משולבת. המטרה של iARM היא לאפשר לחיילים להבין ולהקרין תוצאות טקטיות בצורה יעילה יותר על ידי מתן גישה חלקה למידע קריטי. המערכת כוללת צג מתקדם עם ראש מותקן, מערכת הפעלה משולבת ומודל שירותי נתונים המעבד ומציג מידע רלוונטי. על ידי שילוב של מציאות מוגברת עם בינה מלאכותית, iRAM מאפשרת לחיילים לקבל החלטות בזמן שהם נשארים ניידים.

ULTRA-Vis (Urban Leader Tactical Response, Awareness, and Visualization):

עוד מערכת דומה שפותחה תחת DARPA, ULTRA-Vis פותחה כמערכת אב טיפוס עבור סביבות שדה קרב עירוני. מטרת המערכת היא לספק לחיילים מודעות מצבית מוגברת, כולל מיקומם של גורמים ידידותיים ועויינים כאחד גם כאשר אלמנטים אלו אינם נראים ישירות. המערכת תומכת בזיהוי מחוות באמצעות כפפה לשליטה אינטראקטיבית ובעיקר מסייעת בהקלת שיתוף הפעולה בין חברי כיתה ובקבלת החלטות שלהם בתרחישי לחימה עירונית במתח גבוה.

1. Carreira, Joao, and Andrew Zisserman. "Quo Vadis, Action Recognition? A New Model and the Kinetics Dataset." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 3198–3207.
2. Cao, Zhe, Tomas Simon, Shih-En Wei, and Yaser Sheikh. "Realtime Multi-person 2D Pose Estimation Using Part Affinity Fields." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 7291–7299.
3. Danelljan, Martin, Gustav Hager, Fahad Shahbaz Khan, and Michael Felsberg. "Accurate Scale Estimation for Robust Visual Tracking." *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 743–752.
4. Feichtenhofer, Christoph, Axel Pinz, and Richard Wildes. "Spatiotemporal Residual Networks for Video Action Recognition." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 3056–3064.
5. Kohlschütter, Christian, and Luc Van Gool. "Video-Based Player Tracking for Sports Analysis." *Computer Vision and Image Understanding*, vol. 115, no. 3, 2011, pp. 341–350.
6. Liu, Wei, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C. Berg. "SSD: Single Shot MultiBox Detector." *European Conference on Computer Vision*, 2016, pp. 21–37.
7. Milan, Anton, Stefan Roth, Konrad Schindler, and Josef Sivic. "Object Tracking: A Benchmark." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 9, 2014, pp. 1910–1923.
8. Ng, Joe Yue-Hei, Matthew Hausknecht, Sudheendra Vijayanarasimhan, Oriol Vinyals, Rajat Monga, and George Toderici. "Beyond Short Snippets: Deep Networks for Video Classification." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 4694–4702.
9. Newell, Alejandro, Kaiyu Yang, and Jia Deng. "Stacked Hourglass Networks for Human Pose Estimation." *European Conference on Computer Vision*, 2016, pp. 484–499.
10. Redmon, Joseph, Santosh Divvala, Ross Girshick, and Ali Farhadi. "You Only Look Once: Unified, Real-Time Object Detection." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 779–788.
11. Simon, Tomas, Hanbyul Joo, Iain Matthews, and Yaser Sheikh. "Hand Keypoint Detection in Single Images Using Multiview Bootstrapping." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 3700–3708.
12. Simonyan, Karen, and Andrew Zisserman. "Two-Stream Convolutional Networks for Action Recognition in Videos." *Advances in Neural Information Processing Systems*, 2014, pp. 568–576.
13. Wang, Limin, Yuanjun Xiong, Dahua Dai, and Jiebo Luo. "Temporal Segment Networks: Towards Deeper Representations for Video Understanding." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 4007–4016.
14. Zhang, Li, Yuan Li, and Ram Nevatia. "Robust Object Tracking with Online Multiple Instance Learning." *IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1801–1808.