

Operator Load Assessment

Een evaluatie van de technische centrale van het LCF

De laatste twintig jaar hebben grote veranderingen gebracht in de informatietechnologie. Aan boord van marineschepen zorgt de steeds verdergaande automatisering voor een ingrijpende wijziging in de functie van de operator op de brug, in de commandocentrale en in de technische centrale.

Probleemgebieden of nieuwe mogelijkheden?

Bij toepassing van nieuwe technologieën aan boord doet zich voor de operator een aantal probleemgebieden voor, zoals grote fluctuaties in het informatieaanbod en verbreding van het takenpakket, toenemende automatisering van processen en grotere complexiteit van systemen. Tegelijkertijd hebben we te maken met afnemende personeels- en trainingsbudgetten als gevolg van maatschappelijke, economische en/of politieke druk. Tenslotte

zorgt de (internationale) regelgeving voor extra beperkingen tijdens het ontwerp en gebruik van nieuwe systemen.

de vraag is of we nieuwe technologie optimaal benutten in het ontwerp van nieuwe systemen

Als men echter rekening houdt met deze probleemgebieden kunnen de technologische ontwikkelingen van de laatste decennia veel nieuwe mogelijkheden bieden.

Dit artikel beschrijft bestaande en nieuwe technologieën aan de hand van een evaluatie van de Technische Centrale van het Luchtverdedigings- en Commandofregat (LCF). Na het ontwerp en de implementatie van het *Integrated Monitoring & Control System* (IMCS) van de LC-fregatten ontstond de vraag om het systeem van mens en machine te evalueren. Een team, bestaande uit TNO, DMO, het voormalige CZM en de TOKM, zette het project Operator Load Assessment (OLA) op.



Figuur 1: Veranderingen in informatietechnologie brengen de operator in de problemen.



Figuur 2: Evaluatie van een scenario: de sergeant bekijkt het net uitgevoerde scenario.

Na overleg werden er twee hoofddoelstellingen gedefinieerd²:

1. Evaluatie van het mens-machine systeem. Wat is de invloed van het IMCS op de prestatie en belasting van de operators? Hoe benutten we optimaal de nieuwe mogelijkheden die de technologie ons biedt in het ontwerp van nieuwe systemen?
2. Verkennend onderzoek, het testen van nieuwe technieken. Hoe goed zijn nieuwe technieken praktisch toepasbaar? Welke beperkingen komen we hierbij tegen?

Hieronder worden achtereenvolgens besproken de methode van de evaluatie, de evaluatie van het mens-machine systeem en het verkennend onderzoek. Afgesloten wordt met de conclusie.

Methode van evaluatie

Het IMCS van het LCF is uitgerust met zogenaamde adviesfuncties. Deze adviesfuncties kunnen in verschillende modes ondersteuning bieden aan de operator bij optredende problemen, zoals een onverwachte afwijking van de bedrijfsvoering of een storing in een subsysteem. De OLA evaluatie richt zich specifiek op het effect van deze modes op het werkproces.

Tijdens de evaluatie zijn drie verschillende vormen van IMCS operatorondersteuning getest:

1. Zonder adviesfuncties. De operators werken op de 'ouderwetse' manier, voeren zelfstandig alle benodigde procedurestappen uit, eventueel met behulp van een procedureboek.
2. Adviesfunctie in 'bevestigingsmode'. Bij het optreden van een probleem wordt de uit te voeren procedure als advies door het IMCS aangeboden. De operators moeten elke stap in de procedure apart bevestigen.
3. Adviesfunctie in 'automode'. Wanneer er in deze mode een probleem optreedt worden de procedurestappen, indien mogelijk, automatisch uitgevoerd.

Ten behoeve van de evaluatie zijn drie scenario's ontwikkeld, gericht op problemen met de voortstuwing en storingen in de

automatieken van de voorstuwing (VRA). Het eerste scenario is ontworpen om een veel voorkomende zeewachtsituatie in de Technische Centrale (TC) te simuleren: één complex probleem tegelijkertijd met hoge tijdsdruk. Hierbij kan men denken aan een machinekamer probleem, zoals trillingen in een krachtturbine of een oliemist detectie van een diesel.

Het tweede scenario is ontworpen om een optimale werkdruk te simuleren. De operators moeten enkele niet al te moeilijke problemen onder gemiddelde tijdsdruk oplossen, bijvoorbeeld een laag carter niveau van een diesel of een laag niveau in het startlucht systeem van een gasturbine. Het derde scenario is zo opgezet dat de werkdruk zeer hoog wordt. De operators krijgen meerdere complexe problemen te verwerken onder hoge tijdsdruk. Zo komen er combinaties van alarmen binnen, zoals een hoge temperatuur radiaallager, een 'freeze' alarm van de gasturbine en een lensalarm in de machinekamer.³

Deze scenario's zijn varend uitgevoerd op Hr.Ms. De Ruyter (bevestigingsmode), Hr.Ms. Tromp (zonder adviesfuncties) en op Hr.Ms. Evertsen (automode). Per vaart werkten vier zeewachtdivisies mee, bestaande uit een korporaal, een sergeant en twee matrozen. De evaluatie beperkte zich tot de korporaal en sergeant. Om de scenario's zo reëel mogelijk te maken en de veiligheid te garanderen werkten Imtech, TOKM, brugpersoneel en een *mobility team* van het schip mee. In sessies van 4 tot 5 uur werd er uitleg gegeven, een proefscenario en de drie ontworpen scenario's uitgevoerd en werd het geheel geëvalueerd met de korporaal en sergeant.

Om vast te stellen wat de effecten van de adviesfuncties en scenario's zijn zijn veel data verzameld. Ten eerste hebben de korporaal en sergeant na elk scenario hun eigen videobeelden bekeken (figuur 2) en hebben zij vragenlijsten ingevuld. Ten tweede hebben experts de prestatie van elk team beoordeeld.⁴

Ten derde is er met behulp van videomateriaal bepaald met welke taak de operators op elk moment van het scenario bezig zijn. Op deze manier is veel inzicht verkregen op de invloed van de adviesfuncties op de verschillende taken van de operator. Ten vierde is er ook een groot aantal fysiologische maten opgenomen, zoals ademhalingsfrequentie, hartslag, pupildiameter en spraak (zie verderop).

Resultaten van de evaluatie

De verzamelde informatie en de analyse daarvan geven inzicht in de toepassing van het mens-machine systeemontwerp in de TC aan boord van het LCF. In dit hoofdstuk zal een aantal van deze resultaten getoond worden.

Tabel 1 op de volgende pagina toont de resultaten van de prestatie maat voor de verschillende schepen. Deze prestatie maat is een rapportcijfer tussen de 1 en de 10, beoordeeld door experts. Wanneer de operators extra ondersteuning krijgen in de vorm van adviesfuncties, is er een stijging te zien in prestatie. Wat opvalt is dat deze stijging klein maar significant is, en dat de volledig geautomatiseerde mode geen significante stijging laat zien ten opzichte van de bevestigingsmode. →

	Zonder adviesfuncties	Bevestigingsmode	Automode
Prestatie (expert score)	6.27	7.35	7.49

Tabel 1: Prestatie van de teams bij verschillende ondersteuningsvormen

Figuur 3 toont de bezettingstijd in procenten. Dit is de tijd dat de mensen bezig zijn, ten opzichte van de totale beschikbare tijd. Horizontaal zijn de drie scenario's weergegeven en de verschillende niveaus van ondersteuning. In het 'veel voorkomend probleem'-scenario en het 'hoge werkdruk'-scenario is een duidelijke stijging van de bezettingstijd van de korporaal te zien wanneer de oude situatie, met het procedure-boek en geen ondersteuning van het systeem, vergeleken wordt met de situatie met ondersteuning van het systeem van de adviesfuncties (bevestigingsmode en automode).

Tijdens videoanalyse blijkt dat, vergeleken met de oude situatie, op de schepen met ondersteuning van de adviesfuncties de korporaal taken overneemt van de sergeant. Door de voorgeschreven procedures van de adviesfuncties zijn er meer taken aan de korporaal toegewezen. Dit is overigens niet terug te zien in de bezettingstijd van de sergeant, wel in de prestatie van het hele team.

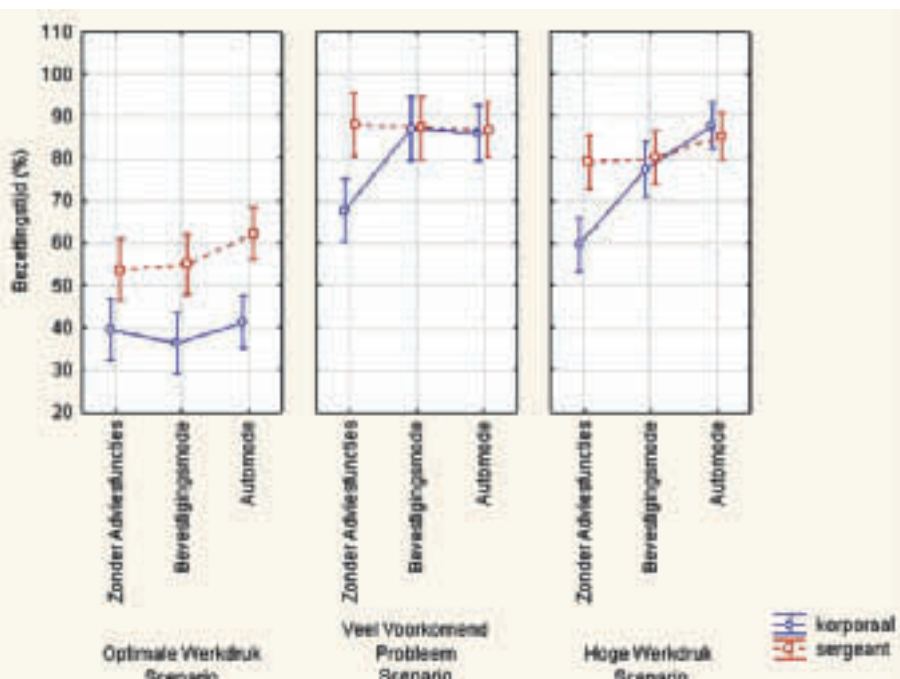
Resultaten verkennend onderzoek

Om in de toekomst de interactie tussen mens en machine te kunnen verbeteren wordt er op dit moment veel verkennend onderzoek gedaan naar adaptieve automatisering. Dit is automatisering die, afhankelijk van bijvoorbeeld het werkaanbod en de belasting van de operator, meer of minder ondersteuning geeft. Zo kan overbelasting worden voorkomen door op dergelijke momenten de mens meer te ondersteunen vanuit het systeem. Door de ondersteuning op tijd weer te reduceren kan onderbelasting en

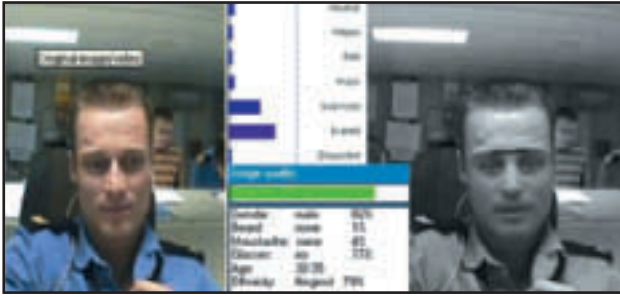
het verlies van vaardigheden worden voorkomen. Zo krijg je als het ware een systeem dat mee beweegt met het werkaanbod. Om dergelijke systemen mogelijk te maken zijn twee bouwstenen essentieel. Ten eerste, het real-time en accuraat kunnen vaststellen van de werkbelasting van een operator. Ten tweede, passende ondersteuningsmodules moeten beschikbaar zijn bij specifieke belastingsproblemen. Een overbelasting door een groot aantal simpele alarmen is bijvoorbeeld iets anders dan overbelasting

adaptieve automatisering kan in de toekomst de interactie tussen mens en machine verbeteren

door een enkel maar zeer complex probleem en vraagt daarom om een andere vorm van ondersteuning. De OLA evaluatie gaf ons de mogelijkheid om ook lange termijn onderzoek uit te voeren met betrekking tot de eerste bouwsteen voor adaptieve automatisering, het real-time vaststellen van werkbelasting in een operationele omgeving. In de evaluatie is daarom gekeken naar de voorspellende waarde en praktische toepasbaarheid van een aantal nieuwe werklust indicatoren. Een van de indicatoren is de gelaatsuitdrukking van operators. Het menselijk gezicht is een uitermate rijk 'informatiedisplay'. Zo zijn mensen vaak in staat een boel informatie af te lezen aan het gezicht van een collega. Ziet een persoon bijvoorbeeld dat een collega op basis van lichaamshouding en gezichtsuitdrukking zeer zwaar belast is, dan kan er gewacht worden met het aanbieden



Figuur 3: De bezettingstijd van de verschillende scenario's afgezet tegen de verschillende ondersteuningsvormen.



Figuur 4: Automatische emotiedetectie aan de hand van gezichtsuitdrukkingen, links het originele beeld, rechts een classificatie "masker" over het gelaat gepresenteerd. (www.noldus.com)

van een nieuw probleem of nieuwe informatie. Momenteel komt commerciële software beschikbaar die tot op zekere hoogte er ook toe in staat is een inschatting te maken van de inspanning op basis van intelligente beeldverwerking. Tijdens de experimenten zijn door middel van kleine camera's tussen de schermen van het werkstation video-opnamen gemaakt. Deze opnamen worden op dit moment geanalyseerd met behulp van software die automatisch het gezicht detecteert en per frame een score op zeven emotie parameters geeft (figuur 4).

Hieruit willen we uiteindelijk een inspanningsscore afleiden. De voorspellende waarde van deze score is op dit moment nog niet beschikbaar. Een praktische beperking is dat de software alleen een score kan geven wanneer de beeldkwaliteit voldoende is. Tijdens OLA bleken de 'rommelige' achtergrond, het vele bewegen van het hoofd en de slechte belichting de beeldkwaliteit te verslechteren. Een groot voordeel is dat de beelden relatief eenvoudig vast te leggen zijn, zonder het operationele proces te verstoren. Ook is real-time verwerking van de data geen probleem.

Een tweede onderzochte indicator is stress in spraak. Op zichzelf is dit geen vreemde maat, omdat een menselijke toehoorder met enige ervaring in de technische centrale eenvoudig het verschil zal kunnen horen tussen onderbelasting en overbelasting. Spraak is ook relatief eenvoudig vast te leggen zonder het operationele proces te verstoren. Door de communicatie in de centrale vast te leggen en deze vervolgens door middel van specifieke software te analyseren kan een real-time indicator voor werklast worden verkregen. Een vereiste is natuurlijk wel dat er sprake is van communicatie, wat een probleem is voor de continuïteit van deze indicator. De eerste resultaten laten zien dat er gedurende het experiment variatie in de emotie en stress kan worden gemeten. Verdere analyse zal moeten aantonen in hoeverre dit correleert met de andere maten. Een operationele beperking levert het achtergrond geluid, wat metingen kan verstoren.

Naast deze meer onbekende maten hebben we de meer klassieke fysiologische maten vastgelegd zoals hartslag, hartslagvariabiliteit en ademhaling. Hierbij was het vooral de vraag in hoeverre deze maten in de operationele context goed zijn vast te leggen. Het veelvuldig bewegen van de operators gedurende de taakuitvoering bleek storend voor zowel de werking van de sensoren als

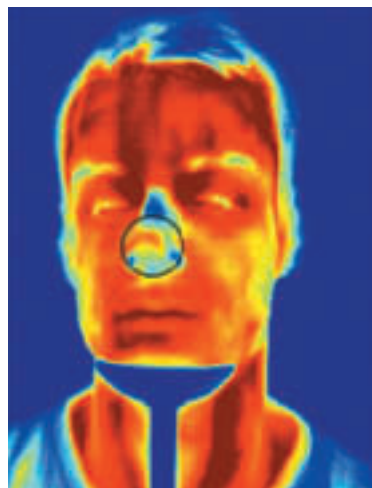
voor de fysiologische toestand zelf. Voor alle genoemde maten geldt dat er verschillen zijn tussen de personen waardoor grote hoeveelheden data vereist zijn om goede voorspellingen te kunnen doen. Daarnaast werken de meeste fysiologische maten met enige vertraging, ze reageren altijd achteraf op bijvoorbeeld overbelasting. Aanpassingen in ondersteuning komen dan te laat. Een maat die in deze evaluatie niet is meegenomen maar in deze context wel relevant is, betreft de neustemperatuur (zie figuur 5). Eerder onderzoek heeft aangetoond dat de temperatuur van de neus afneemt naarmate de inspanning stijgt. De hiervoor benodigde IR-camera is op dit moment nog te groot om eenvoudig in een werkstation in te bouwen, maar de beperkte impact op het operationeel functioneren en de mogelijkheid om de belasting real-time vast te stellen maken het een veelbelovende maat voor de toekomst.

Conclusie

Dit onderzoek laat zien dat een evaluatie in een operationele omgeving zeer nuttige resultaten kan opleveren die het werk aan boord kunnen optimaliseren. Daarnaast levert het kennis op voor nieuwe toepassingen en nieuwe mogelijkheden. Met optimaliseren bedoelen we niet alleen dat er efficiënter of effectiever gewerkt wordt, maar ook het aangenamer maken van de totale werkomgeving, het geheel van mens en machine, speelt een belangrijke rol.

De integratie van mens en (technische) systeem houdt niet op bij het definiëren en ontwerpen van het systeem alleen. De probleemgebieden uit de inleiding zullen in de toekomst een steeds grotere rol gaan spelen. Om nieuwe technieken in het ontwerp optimaal toe te passen zijn meerdere gebruikersevaluaties tijdens de ontwerpfasen en na de implementatie vereist. Om bij toekomstige schepen te voldoen aan voorgeschreven eisen, zal de gebruiker in het middelpunt van het ontwerp moeten staan en neemt dus het belang van soortgelijke evaluaties toe.

Het ontwerp en de implementatie van de adviesfuncties heeft voor een stijging in prestatie gezorgd bij het TC team. Deze stijging komt voor een deel door de veranderde taak van de operator. Zowel met de adviesfuncties als in de volledig geautomatiseerde modus is de korporaal meer tijd bezig ten opzichte van



Figuur 5: Afbeelding van een IR beeld ter bepaling van de neustemperatuur. Afkoeling van de neustemperatuur is een indicatie voor een toegenomen inspanning.

het werken met het procedureboek. Het blijkt dat door het gebruik van adviesfuncties de korporaal extra taken toebedeeld krijgt. Onafhankelijk van het oordeel of dit een positief of negatief (bij)effect is van de adviesfuncties, de verschuiving van taken en verantwoordelijkheden kan grote gevolgen hebben binnen de organisatie. Deze conclusie toont tevens het belang aan van een ontwerptraject waarin scenario's en gebruikers vroegtijdig betrokken worden, waardoor dergelijke gevolgen van systeemontwerp naar voren komen. Of de stijging in prestatie opweegt tegen de extra besteedde tijd verdient nader onderzoek en discussie. Ook het kleine verschil tussen de bevestigings- en automode verdient nader onderzoek. Hierbij dient rekening gehouden te worden met de beperkingen op het moment van het experiment. Zo was er nog weinig ervaring met de adviesfuncties aan boord en waren niet alle IMCS-functionaliteiten geïmplementeerd. Prestatie bij het gebruik van adviesfuncties, in bevestigings- en automode, kan hierdoor in de toekomst verder toenemen.

De resultaten hebben ons tevens inzicht gegeven in de bruikbaarheid en operationele beperkingen van nieuwe meetmethoden en data. Om een uitspraak te kunnen doen over de werkbelasting van de operator lijkt een *combinatie* van meerdere soorten data, zoals spraak, gelaatsuitdrukkingen en neustemperatuur een goede volgende stap. Gekoppeld aan beschikbare context informatie, bijvoorbeeld het aantal of type actieve alarmen, zou de ondersteuning dan aangepast kunnen worden aan de hand van de werksituatie.

Met dank aan de inspanningen van het projectteam en drie enthousiaste bemanningen kunnen we terugkijken op een succes-

volle evaluatie, met bruikbare resultaten voor de korte en lange termijn. ←

LTZT 2 OC ir. M. Grootjen is promovendus Adaptieve Automatisering en Mens Machine Interactie, drs. J. Lindenberg is onderzoeker en programmaleider bij TNO Defensie en Veiligheid, drs. Ing. T.E. de Greef werkt als wetenschappelijk onderzoeker Mens Machine Interactie bij TNO Defensie en Veiligheid. Prof. dr. M. A. Neerinx is hoofd van de Intelligente User Interface groep, TNO Defensie en Veiligheid. Ook is hij als professor Mens Machine Interactie verbonden aan de Faculteit Elektrotechniek, Wiskunde en Informatica aan de Technische Universiteit Delft.

Noten

1. Voor de inhoud van dit artikel zijn deze probleemgebieden minder relevant. Als u hierin geïnteresseerd bent kunt u contact opnemen met de auteurs (marc@grootjen.nl).
2. Naast deze doelstellingen en vragen zijn ook andere onderwerpen onderzocht (bijvoorbeeld: wat kan er in de TC op het LCF beter en wat kunnen we leren voor de volgende generatie schepen), echter deze onderwerpen worden niet in dit artikel behandeld.
3. De scenario's zijn ontwikkeld met behulp van het model voor cognitieve taakbelasting. Neerinx, M. A. 2003. Cognitive task load design: model, methods and examples. In Hollnagel, E. (Ed.), *Handbook of Cognitive Task Design*: 283-305. Mahwah, NJ: Lawrence Erlbaum Associates.
4. Bij een statistische toets gaat men er vanuit dat onderlinge ervaringsverschillen (in een homogene groep) verdeeld zijn over de verschillende te toetsen groepen. Het heeft wel invloed als de groep niet homogeen is, zoals bijvoorbeeld wanneer een groep aan boord van een schip niet of slecht getraind is met de adviesfuncties in vergelijking met een ander schip.

